



A Scoping Review of Voice-Related Patient-Reported Outcome Measures in Korea

Jeong Min Lee ^{ORCID}, Soo Yeon Jung ^{ORCID}, and Han Su Kim ^{ORCID}

Department of Otorhinolaryngology-Head and Neck Surgery, Ewha Womans University College of Medicine, Seoul, Korea

국내에서 개발된 음성장애 환자의 자기보고식 평가도구에 대한 문헌 고찰

이정민 · 정수연 · 김한수

이화여자대학교 의과대학 이비인후과학교실

Received July 25, 2022
Revised September 8, 2022
Accepted September 19, 2022
Address for correspondence
Han Su Kim, MD, PhD
Department of Otorhinolaryngology-Head and Neck Surgery, Ewha Womans University College of Medicine, 1071 Anyangcheon-ro, Yangcheon-gu, Seoul 07985, Korea
Tel +82-2-2650-2686
Fax +82-2-2648-5604
E-mail sevent@ewha.ac.kr

Background and Objectives The purpose of this study is to provide a resource for clinicians and researchers regarding voice-related patient-reported outcome measures (PROMs) available in the Korean Language.

Subjects and Method A systematic search for Korean voice-related PROMs was conducted in December 2021. The search used five electronic databases, including PubMed, Cumulative Index of Nursing and Allied Health Literature, Science Direct, Research Information Sharing Service, DataBase and Periodical Information Academic. Furthermore, reference lists from PROMs-related articles and voice-related journals were independently searched. Search terms emphasized “voice disorders,” “patient-reported,” and “instruments” both in English and Korean. Three independent investigators screened the electronic search results for relevance and selected articles based on the study’s inclusion and exclusion criteria.

Results A total of 164 studies assessing Korean voice-related PROMs were identified, and this list was narrowed to 14 PROMs based on the predetermined inclusion and exclusion criteria. Fourteen voice-related PROMs validated in Korean were selected for a comprehensive review and were categorized as general voice handicap measures, symptom-oriented measures, socioemotional measures, and quality of life measures. Additionally, the characteristics of each of the fourteen Korean voice-related PROMs were summarized in evidence tables.

Conclusion A total of fourteen Korean voice-related PROMs are available for clinical and research use; however, care must be taken when clinicians and researchers consider which instrument(s) to use. The current review provides a useful summary and could serve as a guide for future systematic reviews on Korean voice-related PROMs.

Korean J Otorhinolaryngol-Head Neck Surg 2022;65(11):655-68

Keywords Patient reported outcome measures; Quality of life; Review; Voice disorders.

서론

목소리에 대한 종합적인 평가는 음성장애를 정확하게 진단하고 치료하기 위해 반드시 시행되는 과정이다. 이를 위해 시행되는 객관적인 평가 방법으로는 후두 내시경 검사와 같은

신체검사 평가, 임상가가 환자의 목소리를 듣고 음질(voice quality)을 판단하는 청 지각적 평가, 그리고 음향 분석 프로그램을 활용한 음향학적 평가 방법이 있다.^{1,2)} 이러한 객관적인 평가 방법들은 음성장애 정도를 계량화 할 수 있어 음성장애 환자의 치료 전과 치료 후의 변화를 추적하는 데 용이하다는 장점이 있다.²⁾ 그러나 객관적으로 평가된 음성 질환의 정도가 유사한 경우라 할지라도 개별 음성장애 환자가 처한 직업 상황과 사회활동 정도, 즉 목소리가 직업 또는 사회활동

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<https://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

에서 중요하게 사용되는 정도에 따라 환자가 느끼는 불편감은 다르게 인식될 수 있다.¹⁻³⁾ 따라서 객관적인 평가 방법만으로는 음성장애 환자들이 서로 다르게 인식하고 있는 주관적인 불편감을 평가하는 데 어려움이 따를 수 있다.^{2,3)} 이처럼 음성장애로 진단되어 치료받은 환자들 간의 차이점을 평가하고, 질병 회복 과정에서 환자가 겪는 경험을 더욱 정확하게 측정하기 위해 음성장애 환자의 자기보고식 평가도구들(patient-reported outcome measures)이 꾸준히 개발되어 왔다. 국내에서도 우리나라 음성장애 환자의 실태를 반영한 다양한 설문도구의 번안과 개발이 2007년부터 이루어져 왔고, 개발된 도구에 대한 신뢰도 및 타당도 검증이 활발히 이루어지고 있다.^{1,2)}

본 연구에서는 관심 주제에 대한 정보를 수집하고 평가하기 위해 수행하는 주제범위 문헌고찰 방법(scoping reviews)⁴⁾을 사용하여 국내 음성장애 환자를 위해 번안·개발된 자기보고식 평가도구의 특성을 파악하고, 각 평가도구들의 주요 측정 개념과 임상 활용성(clinical utility)을 비교·분석한 결과를 제시함으로써 진료 현장에서 사용할 수 있는 적절한 평가도구를 선정하는 데 필요한 정보를 제공하고자 한다.

대상 및 방법

연구 대상 및 자료수집

본 연구는 Arksey와 O'Malley⁴⁾가 제시한 주제범위 문헌고찰 방법을 사용하여 연구 절차를 수행하였다. 문헌 검색을 위하여 국외 데이터베이스(PubMed, Cumulative Index to Nursing and Allied Health Literature, Science Direct)와 국내 데이터베이스(Research Information Sharing Service, RISS; DataBase and Periodical Information Academic, DBPIA)를 활용하였다. 검색어는 Pyo⁵⁾와 Francis 등⁶⁾의 선행연구에서 사용된 검색 용어를 참고로 하여 의학주제표목(Medical Subject Headings: MeSH)을 중심으로 구성하였다. 국외 데이터베이스 검색 시 문헌검색을 위해 문헌 제목과 초록을 기준으로 음성장애(voice disorder, voice handicap, vocal disorder, vocal handicap, vocal disability), 환자 자기보고(patient-reported), 평가도구·척도(outcome measure, index, scale), 설문지(survey, questionnaire, instrument), 심리측정 속성(psychometric)을 포함하는 단어들을 불리언 연산자 OR로 연결하여 검색하였다. 검색한 결과를 토대로 '한국(Korea), 한국어(Korean)'를 불리언 연산자 AND를 적용하여 다시 검색하여 문헌을 추출하였다.

국내 데이터 베이스를 사용한 문헌 검색은 '음성장애' 관련 단어들을 ('음성장애' OR '발성장애' OR '음성' OR '발성') 우

선 검색한 결과를 토대로 '지수', '주관적 평가', '삶의 질' 단어를 활용하여 결과 내 재검색을 하였다. 이와 더불어 국내외 음성장애 관련 학술지(Journal of Voice, Communication Sciences and Disorders, Korean Journal of Otorhinolaryngology-Head and Neck Surgery, Journal of the Korean Society of Laryngology, Phoniatrics and Logopedics)들을 대상으로 세부 검색을 하였고, 검색된 평가도구 문헌의 참고 문헌 및 해당 문헌을 인용한 목록을 확인하여 누락되는 도구가 없도록 하였다. 국내에서 출판된 문헌을 포괄적으로 검색하기 위하여 문헌의 출판일은 영미권에서 최초로 개발된 음성장애 관련 자기보고식 평가도구의 출판이 1984년도에 이루어졌다는 Francis 등⁶⁾의 보고를 참고하여 1980년 1월 1일부터 2021년 12월 20일까지로 설정하였다. 문헌 검색 시, 언어는 영문 또는 국문으로 출판된 문헌으로, 종류는 학술지에 게재된 논문과 학위 논문으로 한정하였다. 자료 수집을 위한 문헌검색은 2021년 12월 15일부터 12월 20일까지 수행하였다. 이와 같은 과정을 거쳐 검색된 문헌들은 3인의 연구자(이비인후과 전문의 2인과 언어재활학 박사)가 선정한 기준을 토대로 문헌 선정이 이루어졌다(Table 1).

자료 선정 및 분석

국내·외 데이터베이스를 통해 문헌을 검색한 결과, PubMed 52편, CINAHL 10편, Science Direct 16편, RISS 49편, DBPIA 28편으로 155편의 문헌이 검색되었고 검색된 평가도구들의 참고문헌과 해당 문헌을 인용한 논문을 확인하여 9개의 문헌을 추가로 발견하였으며, 그중 중복된 논문 31편을 제외한 총 133편의 문헌이 일차적으로 선별되었다. 이들 문헌의 제목과 초록을 검토하여 Table 1에 제시된 포함기준에 적절하지 않은 문헌 110편을 제외하였다. 제목 및 초록 검토 후 선별된 23편 논문의 본문을 검토한 결과, 음성장애 환자의 자기보고식 평가도구의 개발과 연관성이 없거나 심리측정적 속성을 검증하지 않은 연구 7편, 예비 연구 2편, 목소리와 관련된 환자 보고가 포함되지 않은 연구 1편을 추가로 제외하여 13개의 문헌이 분석 대상에 포함되었다. 그리고 이들 문헌 중 하나의 문헌에서 두 개의 평가도구에 대한 심리측정 속성을 보고한 경우는 각각을 하나의 연구로 간주하여 최종적으로 14개의 문헌이 선정되었다(Fig. 1).

음성장애 환자를 위해 개발된 자기보고식 평가도구의 체계적인 분석을 위하여 Francis 등⁶⁾에 제시된 분석표를 기초로 평가도구별 특성과 임상실용성을 요약한 분석표를 작성하였다. 평가도구별 특성은 평가도구의 한글명, 국내 출판 연도, 원 도구의 개발자 및 출판연도, 측정 대상 정보, 평가 하위 항목 등의 내용들을 분석하였다(Table 2). 임상 실용성은 평

가내용에 따른 도구 유형, 환자군에 대한 적용(범용) 가능성, 적일치도, 검사-재검사 신뢰도, 타당도를 포함하는 내용들을 평가도구에 사용된 척도 특성, 채점 방식, 심리측정적 요인인 설문도구 개발 당시 참여한 대상자 정보(성별·연령 분포), 내

Table 1. Inclusion and exclusion criteria for included studies

Inclusion	Exclusion
Population	
- Korean patients with voice disorders	
- Professional and/or occupational voice users	
Age	
- Children	
- Adults	
- Elders	
Topic of study	
- Voice problems or hoarseness	- Speech problems
- Patient-reported outcome, instrument, questionnaire, or survey	
Types of studies	
- Cross-cultural translation & validation of a patient-reported outcome measure	- Treatment trials
- Validity and/or reliability testing of developed patient-reported outcome measure	- Case reports and/or series
Type of publication	
- Peer-reviewed publications	- Non peer-reviewed publications
- Published Master's thesis or doctoral dissertation	- Review/opinion articles
	- Conference presentations
	- Published abstracts

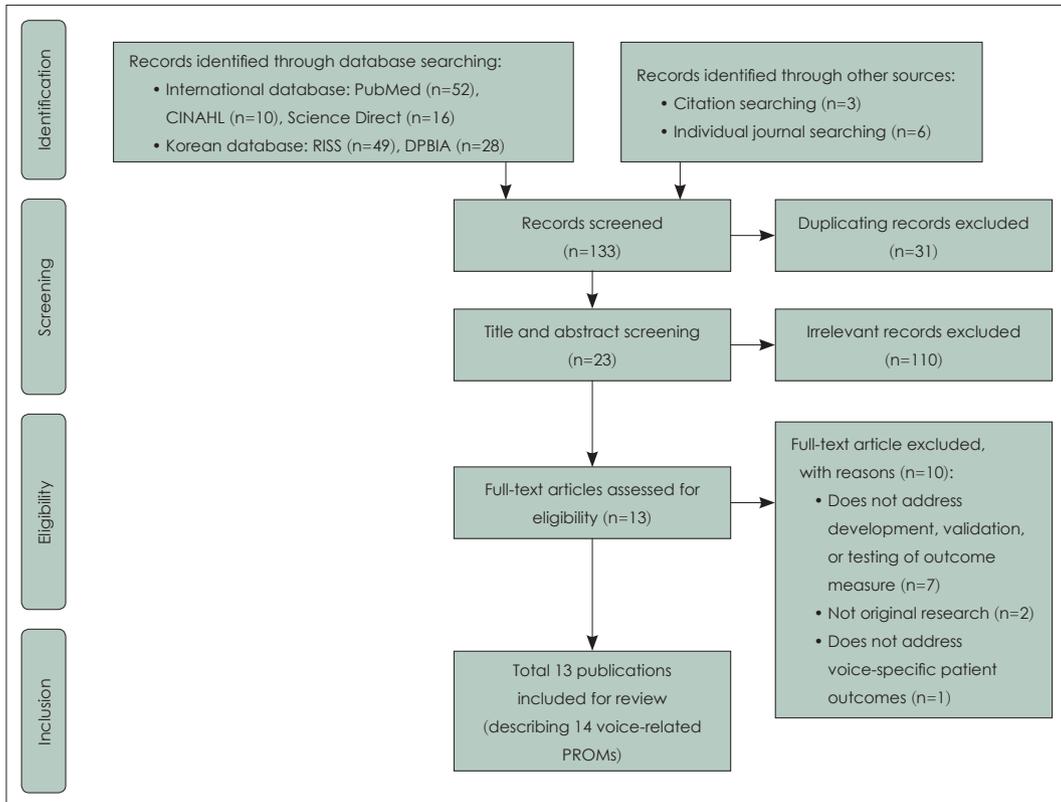


Fig. 1. Study search process with search results. CINAHL, Cumulative Index of Nursing and Allied Health Literature; RISS, Research Information Sharing Service; DPBIA, DataBase and Periodical Information Academic; PROMs, patient-reported outcome measures.

Table 2. Voice-related patient-reported outcome measures (PROMs) developed in Korean (n=14)

Voice-related PROMs in Korean	Article (year)	Original PROMs	Country of origin	Target population	Source of report	Subscales
Korean Version of the Voice-related Quality of Life (K-VQOL; 한국어판 음성관련 삶의 질 척도)	Kim, et al. ² (2007)	Voice-related Quality of Life (VQOL; Hogikyan and Sethuraman, ³⁰ 1999)	USA	Individuals with voice disorders	Self-report	Functional Physical Emotional
Korean Voice Handicap Index (K-VHI; 한국어판 음성장애지수)	Kim, et al. ² (2007) and Yun ¹⁾ (2007)	Voice Handicap Index (VHI; Jacobson, et al., ³ 1997)	USA	Individuals with voice disorders	Self-report	Functional Physical Emotional
Korean Voice Handicap Index-10 (K-VHI-10; 한국어판 음성장애지수-10문항)	Yun ¹⁾ (2007)	Voice Handicap Index-10 (VHI-10; Rosen, et al., ⁸⁾ 2004)	USA	Individuals with voice disorders	Self-report	Functional Physical Emotional
Korean Singing Voice Handicap Index (K-SVHI; 한국어판 가수음성장애지수)	Lee and Sim ¹⁰⁾ (2013)	Singing Voice Handicap Index (sVHI; Cohen, et al., ⁹⁾ 2007)	USA	Singers with voice disorders	Self-report	Physical Social Emotional
Korean Pediatric Voice Handicap Index (K-pVHI; 한국어판 소아음성장애지수)	Park, et al. ¹⁵⁾ (2013)	Pediatric Voice Handicap Index (pVHI; Zur, et al., ¹⁴⁾ 2007)	USA	Children with voice disorders	Parental/ caregiver report	Economic Functional Physical Emotional
Korean Children's Voice Handicap Index-10 (K-CVHI-10; 한국어판 아동음성장애지수-10문항)	Ra ¹⁷⁾ (2015)	Children's Voice Handicap Index-10 (CVHI-10; Ricci-Maccarini, et al., ¹⁶⁾ 2013)	Italy	Children with voice disorders	Parental/ caregiver and self-report	No subscales
Korean Vocal Fatigue Index (K-VFI; 한국어판 음성피로도 검사)	Kim ²¹⁾ (2016)	Vocal Fatigue Index (VFI; Nanjundeswaran, et al., ²⁰⁾ 2015)	USA	Individuals with voice disorders	Self-report	Tiredness Physical Improvement with rest
Korean Voice Activity and Participation Profile (K-VAPP; 한국어판 음성활동 및 참여 프로파일)	Lee, et al. ²⁷⁾ (2016)	Voice Activity and Participation Profile (VAPP; Ma and Yiu, ²⁶⁾ 2001)	Hong Kong	Individuals with voice disorders	Self-report	*Composite scores: Activity Limitation, Participation Restriction *Subscales: Self-perceived severity, Job, Daily communication, Social communication, Emotion
Korean Voice Symptom Scale (K-VoISS; 한국어판 음성증상척도)	Son, et al. ¹⁹⁾ (2018)	Voice Symptom Scale (VoISS; Deary, et al., ¹⁸⁾ 2003)	Great Britain	Individuals with voice disorders	Self-report	Functional Physical Emotional

결 과

평가도구별 분석

한국어판 음성장애지수(Korean Voice Handicap Index, K-VHI)

Voice Handicap Index (VHI)는 가장 대표적이고 널리 사용되는 음성장애와 관련된 자가보고식 평가도구로서, 성인 음성장애 환자가 일상에서 느끼는 불편 정도를 측정하기 위해 1997년 Jacobson 등³⁾에 의해 최초로 개발되었다. 국내에서는 한국 음성장애 환자들의 정서에 맞고, 이해하기 쉬운 표현으로 번역된 두 번안본이 Kim 등²⁾과 Yun¹⁾에 의해 2007년 개발되었으며 신뢰도와 타당도가 검증되었다. VHI는 총 30개의 문항과 3개의 하부영역으로 구성되어 기능, 신체, 정서적 측면에서 목소리와 관련된 불편 정도를 평가한다.^{1,3)} 각 문항은 5점 리커트(Likert) 척도(0=“전혀 그렇지 않다” ~4=“항상 그렇다”)를 사용하여 평가되며 합산점수는 0점에서 120점까지의 범위를 보이며, 점수가 높을수록 음성장애로 인한 일상의 불편이 큰 것을 의미한다. VHI는 현재 20개 이상의 언어로 번역되어 임상과 연구 현장에서 가장 널리 사용되고 있는 음성장애 관련 자기보고식 설문도구이다.⁵⁾ 이와 더불어 VHI는 후두암, 근긴장성 발성장애, 연축성 발성장애 등을 아우르는 다양한 환자들에게 적용되었으며, 음성장애 치료 후 일어나는 변화를 단·장기적으로 추적하는데 국내·외에서 가장 많이 사용되는 도구이다.^{1,2,5)} 한국어판 VHI의 절단점(cut-off score)은 Lee와 Kim⁷⁾에 의해 12.1점으로 보고되었으며, 총점이 해당 점수 이상일 경우 음성장애로 인한 일상에서의 불편감이 큰 것으로 해석할 수 있다.

한국어판 음성장애지수-10(Korean Voice Handicap Index-10, K-VHI-10)

2004년 Rosen 등⁸⁾은 설문도구를 작성하는데 걸리는 시간적 부담을 최소화하여 임상에서 사용이 용이하도록 VHI의 문항을 10개로 줄인 Voice Handicap Index-10 (VHI-10)을 개발하였다.¹⁾ Rosen 등⁸⁾은 VHI의 30개 문항 중, 음성장애 환자군과 정상 대조군 간에 가장 차이를 많이 보인 문항과 총 6명으로 구성된 음성장애 관련 전문가 패널들이 추천한 문항들을 바탕으로 VHI-10을 개발하였으며, 이후 신뢰도와 타당도를 검증하였다.¹⁾ 국내에서는 Yun¹⁾에 의해 K-VHI의 30개 문항들을 토대로 K-VHI-10이 개발되었으며, 후두암을 제외한 다양한 음성장애 환자들을 대상으로 신뢰도와 타당도를 검증하여 국내에서의 유용성이 입증되었다. VHI-10은 기존 VHI의 하부영역을 대표하는 문항들(기능 5문항, 신체 3문항,

Table 2. Voice-related patient-reported outcome measures (PROMs) developed in Korean (n=14) (continued)

Voice-related PROMs in Korean	Article (year)	Original PROMs	Country of origin	Target population	Source of report	Subscales
Korean Voice Catastrophization Index (K-VCI; 한국어판 음성 파국화 지수)	Lee, et al. ²⁹⁾ (2019)	Voice Catastrophization Index (VCI; Shoffell-Havakuk, ²⁸⁾ 2019)	USA	Individuals with voice disorders	Self-report	Rumination Magnification Helplessness
Korean Aging Voice Index (K-AVI; 한국어판 노인음성지수)	Bae, et al. ¹²⁾ (2019)	Aging Voice Index (AVI; Etter, et al., ¹¹⁾ 2019)	USA	Elders with voice disorders	Self-report	No subscales
Aging Voice Index-Korean (AVI-KR; 한국어판 노인성음성지수)	Pyo, et al. ¹³⁾ (2021)	Aging Voice Index (AVI; Etter, et al., ¹¹⁾ 2019)	USA	Elders with voice disorders	Self-report	No subscales
Korean Vocal Tract Discomfort Scale (K-VTDS; 한국어판 성도불편감척도)	Lee, et al. ²³⁾ (2022)	Vocal Tract Discomfort Scale (VTDS; Mathieson, ²³⁾ 1993)	Great Britain	Individuals with voice disorders	Self-report	Frequency of sensation/ symptom Severity of sensation/ symptom
Thyroidectomy-related Voice and Symptom Questionnaire (TVSQ; 갑상선 수술 관련 음성 및 증상 설문)	Hwang, et al. ²⁵⁾ (2022)	Thyroidectomy Related Voice Questionnaire (TVQ; Nam, et al., ²⁴⁾ 2012)	Korea	Patients undergoing thyroidectomy	Self-report	Voice change Throat and neck discomfort

Table 3. Comparison of practical features between voice-related PROMs in Korean (n=14)

Type of PROM	Versatility				Practicality			Psychometric features						
	Use in patients with diverse diagnostic categories of dysphonia	Use in specific age groups or professional voice users	Use in patients undergoing thyroid surgery	Number of items	Administration time	Response method	Range of total score	Scoring method	Total score cutoff	Participants included in the PROM validation	Internal consistency reliability ($\alpha > 0.90$)	Test-retest reliability (> 0.90)	Validity	
General Voice Handicap measures														
K-VHI	+		+	30	-	5-point Likert scale (0-4)	0 (best)-120 (worst)	Sum of items	12	Cases=156 (68:88) VNCs=66 (33:33)	Cases=45 (13), 20-80 VNCs=45.4 (15.5), 20-75	0.97	0.97	Concurrent, Construct, Factorial
K-VHI-10	+		+	10	-	5-point Likert scale (0-4)	0 (best)-40 (worst)	Sum of items	-			0.92	0.97	Concurrent, Construct, Factorial
K-SVHI		+		36	-	5-point Likert scale (0-4)	0 (best)-144 (worst)	Sum of items	-	Cases=40 VNCs=90	20 years and older	0.95	0.99	Concurrent, Construct
K-AVI		+		23	-	5-point Likert scale (0-4)	0 (best)-92 (worst)	Sum of items	-	Cases=111 (57:44) VNCs=100 (55:45)	Cases=69.8, 65-80 VNCs=70.6, 65-82	0.83	0.85	Content, Concurrent, Construct
AVI-KR									11	Cases=76 (38:38) VNCs=50 (17:33)	Cases=69.62 (6.95), 60-87 VNCs=65.38 (2.49), 60-85	0.98	0.97	Content, Concurrent, Construct
K-pVHI		+		23	-	5-point Likert scale (0-4)	0 (best)-92 (worst)	Sum of items	-	Cases=41 (30:11) VNCs=60 (29:31)	Cases=8.95 (1.94), 7-12 VNCs=9.5 (1.72), 7-12	0.92	0.97	Content, Construct
K-CVHI-10		+		10	-	5-point Likert scale (0-4)	0 (best)-40 (worst)	Sum of items	3.5	Cases=100 (46:54) VNCs=100 (44:56)	Cases=10.14 (0.84), 7-12 VNCs=8.83 (1.64), 7-12	0.91 for children	0.76 for children	Content, Concurrent, Construct
General Voice Handicap & Symptom-oriented measure														
K-VoISS	+		+	44	-	5-point Likert scale (1-5)	44 (best)-220 (worst)	Sum of items	-	Cases=31 (6:25)	45.8 (10.7)	0.97	>0.90	Content, Concurrent, Construct

Table 3. Comparison of practical features between voice-related PROMs in Korean (n=14) (continued)

Type of PROM	Versatility			Practicality			Psychometric features							
	Use in patients with diverse diagnostic categories of dysphonia	Use in specific age groups or professional voice users	Use in patients undergoing thyroid surgery	Number of items	Administration time	Response method	Range of total score	Scoring method	Total score cutoff	Participants included in the PROM validation	Internal consistency reliability ($\alpha > 0.90$)	Test-retest reliability (> 0.90)	Validity	
Symptom-oriented measures														
K-VFI	+			19	-	5-point Likert scale (0-4)	0 (best)-76 (worst)	Sum of items	-	SLPs=20 Preschool teachers=30 VNCs=29 (all female)	SLPs=34.8 (7.79) Preschool teachers=29 (7.53) VNCs=30.1 (5.6)	-	-	Concurrent, Construct
K-VTDS	+			8	-	7-point Likert scale (0-6)	0 (best)-48 (worst)	Sum of items	-	Cases=159 (79:80) VNCs=131 (56:75)	Cases=48.6 (14.8), 20-80 VNCs=47.7 (14.2), 20-75	0.91	0.93	Concurrent, Construct
TVSQ		+		20	-	5-point Likert scale (0-4)	0 (best)-80 (worst)	Sum of items	-	Cases=50 VDCs=50 VNCs=30	Not specified	-	0.76	Content, Face, Concurrent, Construct
Socio-emotional measures														
K-VAPP	+			28	-	11-point EAI scale (0-10)	0 (best)-280 (worst)	Sum of items and score standardization	14.5	Cases=150 (42:108) VNCs=150 (42:108)	Cases=45.15 (15.15), 19-86 VNCs=42.31 (11.86), 20-71	0.97	0.96	Content, Concurrent, Construct
K-VCI	+			13	-	5-point Likert scale (0-4)	0 (best)-65 (worst)	Sum of items	6	Cases=80 (39:41) VNCs=25 (12:13)	Cases=57.4 (10.4) VNCs=55.8 (11.5)	-	0.79-0.93	Face, Concurrent, Construct
Quality of Life measure														
K-VQOL	+			10	-	5-point Likert scale (1-5)	0 (worst)-100 (best)	Score standardization	91	Cases=131 (54:77) VNCs=111 (43:68)	Cases=33.73, 15-49 VNCs=29.68, 20-46	0.96	0.90	Concurrent, Construct

PROMs, patient-reported outcome measures; M, mean; SD, standard deviation; VNCs, vocally normal controls; VDCs, voice disordered controls; SLPs, speech language pathologists; K-VHI, Korean Voice Handicap Index; K-VHI-10, Korean Voice Handicap Index-10; K-SVHI, Korean Singing Handicap Index; K-AVI, Korean Aging Voice Index; AVI-KR, Aging Voice Index-Korean Revised; K-pVHI, Korean Pediatric Voice Handicap Index; K-CVHI-10, Korean Children's Voice Handicap Index-10; K-VoIS, Korean Voice Symptom Scale; K-VFI, Korean Vocal Fatigue Index; K-VTDS, Korean Vocal Tract Discomfort Scale; TVSQ, Thyroidectomy-related Voice and Symptom Questionnaire; K-VAPP, Korean Voice Activity and Participation Profile; K-VCI, Korean Voice Catastrophization Index; K-VQOL, Korean Voice-related Quality of Life

정서 2문항)로 구성되어 있으며, 각 문항은 VHI와 마찬가지로 5점 리커트 척도를 사용하여 평가한다.^{1,2)} 합산점수는 0점에서 40점까지의 범위를 보이며, 점수가 높을수록 화자가 음성장애로 인해 경험하는 일상의 불편이 큰 것으로 해석할 수 있다.^{1,2)} VHI-10은 약 19개 국가의 언어로 번안되어 VHI 다음으로 가장 많은 나라에서 사용되는 음성장애 관련 자기보고식 평가도구이다.⁵⁾ 아직 국내 음성장애 환자들을 위한 K-VHI-10의 절단점은 보고되지 않았으나, VHI-10은 문항 수가 적고 작성이 간편하여 국내외 임상과 연구 현장에서 활발히 사용되고 있다.¹⁾

한국어판 가수음성장애지수(Korean Singing Voice Handicap Index, K-SVHI)

Singing Voice Handicap Index (SVHI)는 기존에 개발된 VHI가 가수들이 경험하고 있는 목소리 관련 불편감을 충실히 반영하지 못한다는 제한점을 극복하기 위해 2007년 Cohen 등⁹⁾에 의해 개발되었다. 국내에서는 2013년 Lee와 Sim¹⁰⁾에 의해 음성장애로 진단받은 가수들 40명과 정상 음성을 가진 가수들 90명을 대상으로 한국어판 SVHI가 번안되고 신뢰도와 타당도가 검증되었으며, 소프라노 성악가, 실용음악 보컬 전공 대학생 등을 대상으로 K-SVHI가 활용되었다. SVHI는 총 36개의 문항과 4개의 하부영역(신체, 정서, 사회, 경제적 측면)으로 구성되어 있으며, 각 문항은 0-4점을 사용하여 평가된다.^{9,10)} 총점수는 0점에서 144점까지의 범위를 보이며, 절단점은 보고된 바 없으나 총점이 높을수록 가수가 현재 경험하고 있는 목소리로 인한 불편이 높은 것으로 해석할 수 있다.¹⁰⁾

한국어판 노인(성)음성지수(Korean-Aging Voice Index, Aging Voice Index - Korean Revised; K-AVI, AVI-KR)

Aging Voice Index (AVI)는 2019년 Etter 등¹¹⁾이 60세 이상의 고령인구를 대상으로 목소리로 인한 불편감이 삶의 질에 미치는 영향을 평가하기 위해 개발한 설문도구이다. 국내에서는 한국 노인들의 정서에 맞고, 이해하기 쉬운 표현으로 번역된 한국어판 노인(성)음성지수(K-AVI, AVI-KR)가 Bae 등¹²⁾과 Pyo 등¹³⁾에 의해 각각 개발되었다. 두 연구 모두 동일하게 AVI를 한국어로 번안하고 국내 음성장애 노년층 환자를 대상으로 검증하였으나, 설문도구의 개발 과정에서 차이가 있다. Bae 등¹²⁾에 의해 개발된 K-AVI (한국어판 노인음성지수)는 음성장애로 진단받거나 정상 음성을 가진 만 65세 이상의 노인을 대상으로 타당도와 신뢰도를 검증하였으나, 이와 달리 Pyo 등¹³⁾에 의해 개발된 AVI-KR (한국어판 노인성음성지수)은 음성장애로 진단받거나 정상 목소리를 가졌으며, 목소리 산출에 영향을 줄 수 있는 약물을 복용하지 않은 60세 이상

의 노인을 대상으로 설문도구를 검증하였다. 또한 Pyo 등¹³⁾은 노년층의 성별에 따른 목소리 관련 불편감을 추가로 보고하였다. 이와 더불어 두 연구는 음성장애 환자군 구성에서 차이가 보고되었는데, 예를 들어 Pyo 등¹³⁾은 양성 성대질환(benign vocal fold lesion)을 겪고 있는 노인들을 Bae 등¹²⁾의 연구보다 더 많이 포함하여 AVI-KR의 신뢰도와 타당도를 검증하였다. K-AVI와 AVI-KR은 번안·개발 과정에서 약간의 차이가 보고되었으나, 두 번안본 모두 국내 노년층을 대상으로 AVI의 유용성을 입증하였다는 데 의의가 있다.

AVI는 총 23문항으로 이루어져 있으며, 1-18번 문항은 0-4점으로 평가하고, 19-23번 문항은 4-0점으로 평가하게 된다.¹³⁾ 합산 점수는 0점에서 92점까지의 범위를 보이며, 점수가 높을수록 노인 환자가 음성장애로 인해 경험하는 일상에서의 불편이 큰 것을 의미한다.¹¹⁻¹³⁾ Pyo 등¹³⁾은 국내 노년층을 위한 AVI-KR의 총점이 절단점인 11점을 넘어갈 때 음성장애가 노인 환자의 일상생활에 미치는 영향이 큰 것으로 해석할 수 있음을 보고하였다.

한국어판 소아음성장애지수(Korean Pediatric Voice Handicap Index, K-pVHI)

Pediatric Voice Handicap Index (pVHI)는 2007년 Zur 등¹⁴⁾이 아동의 음성장애 정도를 평가하기 위해 기존에 성인용으로 개발된 VHI를 부모보고 형식으로 변경하여 개발한 평가도구이다. pVHI는 아동이 인지하고 있는 목소리 관련 불편감을 기능, 신체, 감정적 어려움의 영역으로 나누어 평가하도록 구성되어 있으며, 16개 이상의 언어로 번안되어 아동을 위한 대표적인 음성장애 설문도구로서 임상과 연구 현장에서 널리 활용되고 있다.¹⁵⁾ 국내에서는 2013년 Park 등¹⁵⁾에 의해 한국어판이 개발되었으며, 성대 결절로 진단받은 아동 41명과 정상 음성을 가진 아동 60명을 대상으로 신뢰도와 타당도가 검증되어 연구와 임상 현장에서 사용되고 있다. 해당 도구는 총 23문항으로 이루어져 있으며, 각 문항을 아동의 보호자가 0-4점으로 평가하게 된다. 합산점수는 0점에서 92점까지의 범위를 보이며, 점수가 높을수록 아동이 음성장애로 인해 경험하는 일상에서의 불편이 크다고 해석할 수 있으나 아직 한국어판 pVHI의 절단점은 보고되지 않았다.¹⁵⁾

한국어판 아동음성장애지수-10 (Korean Children's Voice Handicap Index-10, K-CVHI-10)

Children's Voice Handicap Index-10 (CVHI-10)은 Ricci-Maccarini 등¹⁶⁾에 의해 2013년 개발된 10문항인 자기보고식 설문도구로 0-4점을 이용하여 아동과 보호자가 모두 목소리와 관련된 일상에서의 불편을 보고할 수 있도록 구성

되어 있다. 해당 평가도구는 개발 당시 설문 문항을 아동이 읽었을 때 문항을 적절히 이해하고 답변할 수 있도록 내용과 표현들을 아동의 입장에서 검토하여 만들어졌다. 국내에서는 2015년 Ra¹⁷⁾에 의해 음성장애로 진단받은 아동 100명, 정상 음성인 아동 100명을 대상으로 한국어판 CVHI-10이 번안되고 신뢰도와 타당도가 검증되었다. K-CVHI-10의 절단점은 3.5점으로 보고되었으며, 해당 도구의 총점이 절단점을 넘어갈 때 음성장애가 아동의 일상생활에 미치는 영향이 큰 것으로 해석할 수 있다.¹⁷⁾ K-CVHI-10은 아동이 경험하고 있는 목소리와 관련된 일상에서의 불편감을 부모뿐 아니라 아동의 관점에서 측정할 수 있는 유일한 평가도구라는 점에서 의의를 지닌다.^{16,17)}

한국어판 음성증상척도(Korean Voice Symptom Scale, K-VoiSS)

Voice Symptom Scale (VoiSS)은 Deary 등¹⁸⁾이 2003년 음성장애 환자들이 경험하는 전반적인 목소리 관련 불편감과 목소리 관련 증상들을 포괄적으로 측정하기 위해 개발한 도구이다. Deary 등¹⁸⁾은 개방형 질문을 통해 음성장애 환자들에게서 목소리 증상에 관한 답변들을 수집한 뒤, 그중 추려진 31개의 문항과 기존 VHI에서 추출한 13문항을 합치는 과정을 통해 44문항의 VoiSS를 개발하였고, 이후 신뢰도와 타당도를 검증하였다. 국내에서는 2018년에 Son 등¹⁹⁾에 의해 한국어판 VoiSS가 번안되었으며, 갑상선 수술을 시행 받은 국내 환자들을 대상으로 신뢰도와 타당도가 검증되었다. VoiSS 문항들은 음성장애 환자가 경험할 수 있는 목안의 염증 증상, 가래, 목소리 변화와 같은 다양한 목소리 관련 신체적 증상과 음성장애로 인한 감정적 어려움 등을 기술하고 있으며, 해당 문항은 1-5점으로 평가하게 된다.^{18,19)} 총점은 44점부터 220점까지 범위를 보이며 절단점은 보고된 바 없으나, 점수가 클수록 환자가 경험하고 있는 목소리 관련 증상 불편이 큰 것으로 해석된다.¹⁹⁾ VoiSS는 음성장애 관련 전문가들에 의해 문항이 미리 개발된 다른 설문도구들과 달리, 음성장애 환자와들의 인터뷰를 토대로 문항을 추출하여 가장 '환자 중심'으로 설문 개발이 이루어진 음성장애 관련 자기보고식 평가 도구이다.¹⁸⁾ 또한 VoiSS는 음성장애 환자들이 경험하고 있는 신체적 증상과 음성장애 관련 삶의 질을 많은 문항에 걸쳐 기술하고 있어 음성장애 환자들이 경험하고 있는 전반적인 목소리 관련 불편감을 가장 포괄적으로 평가할 수 있는 도구로 평가받고 있다.¹⁸⁾

한국어판 음성피로도검사(Korean Vocal Fatigue Index, K-VFI)

장시간 음성을 사용한 뒤 경험할 수 있는 목소리의 피로감은 목소리 사용이 많은 직업적 음성사용인들이 주로 호소하는 음성장애 증상 중 하나이다.²⁰⁾ 그러나 음성피로도라는 개념을 구체적으로 정의하고, 음성장애의 진단과 평가 시 활용할 수 있는 설문도구의 부재로 인해 Nanjundeswaran 등²⁰⁾은 Vocal Fatigue Index (VFI)를 2015년도에 개발하였다. 국내에서는 2016년 Kim²¹⁾이 한국어판 VFI를 번안하고 직업적 음성사용인을 대상으로 신뢰도와 타당도를 검증하였다. K-VFI는 총 19문항으로 각 문항당 0-4점으로 평가되며, 첫 번째와 두 번째 영역은 음성 피로의 정도, 음성 사용을 기피하는 정도, 목소리 관련 신체적 불편과 관련된 문항으로 이루어져 있으며, 세 번째 영역은 음성 휴식 후 증상 개선에 관한 문항으로 구성되어 있다.^{20,21)} 총점은 0점부터 76점까지 범위를 보이며 절단점은 보고된 바 없으나, 점수가 클수록 화자가 인식하고 있는 음성피로도가 큰 것으로 해석된다.²¹⁾

한국어판 성도불편감척도(Korean Vocal Tract Discomfort Scale, K-VTDS)

1993년 Mathieson²²⁾은 과기능성 음성장애 환자들이 경험하고 있는 성도와 관련된 불편감의 빈도와 중증도를 측정하기 위하여 Vocal Tract Discomfort scale (VTDS)를 개발하였다. 국내에서는 2022년 Lee 등²³⁾에 의해 한국판 VTDS가 번안되었으며 기능적, 구조적, 신경학적 음성장애로 진단받은 음성장애 환자들을 대상으로 신뢰도와 타당도가 검증되었다. K-VTDS는 음성장애 환자들이 경험하고 있는 주된 목소리 관련 증상들을 묻는 8문항으로 구성되어 있으며 환자들은 0-6점 척도를 사용하여 목소리 관련 증상의 빈도와 심한 정도를 평가하게 된다. K-VTDS의 총점은 0점부터 48점까지이며 절단점은 보고되지 않았으나, 총점이 클수록 음성장애 환자가 경험하고 있는 목소리 증상들이 자주 일어나고 그로 인한 불편이 심한 것으로 해석할 수 있다.^{22,23)}

갑상선 수술 관련 음성 및 증상 설문(Thyroidectomy-related Voice and Symptom Questionnaire, TVSQ)

Thyroidectomy-related Voice Questionnaire (TVQ)는 갑상선 수술을 시행 받은 환자들의 목소리 증상들과 관련 변화를 추적하기 위해 2012년 Nam 등²⁴⁾에 의해 개발된 최초의 설문도구이다. 이후 2022년 Hwang 등²⁵⁾은 갑상선 수술을 시행 받은 환자들이 겪을 수 있는 목소리 증상들을 기술한 문항들의 신뢰도와 타당도를 다시 검증하여 TVSQ를 발표하였다. 총 20개의 TVSQ 문항들은 갑상선 수술로 인해 발생

할 수 있는 목안의 통증, 가래, 구강 건조, 사레 들림, 발화시 숨이차는 증상들과 같은 여러 목소리 관련 증상들로 구성되어 있으며, 각 문항들은 0-4점 척도를 사용하여 평가한다.²⁵⁾ 총점은 0점부터 80점까지이며, 점수가 클수록 갑상선 수술 전과 후로 환자가 경험하고 있는 목소리 관련 증상 불편이 큰 것으로 해석된다.^{24,25)} 현재까지 TVSQ의 절단점은 보고된 바 없으나 2012년 개발된 TVQ의 절단점은 5점으로 확인되었으며, Nam 등²⁴⁾은 갑상선 수술을 받기 전 환자가 보고한 TVQ의 총점이 절단점 이상일 경우 후두 내시경 검사, 음향학적 평가와 같은 정밀검사를 시행하도록 권고하였다. 주로 국외에서 개발된 설문도구를 번안하는 과정을 거쳐 개발된 다른 음성장애 관련 자기보고식 평가도구들과 달리 TVSQ는 국문으로 제작되어 국내 환자들을 대상으로 검증이 이루어진 후, 영문으로 역번안 되어 출판된 평가도구이다.^{24,25)} 또한 해당 도구는 갑상선 수술을 시행 받은 환자들이 수술 전후로 경험할 수 있는 목소리 불편감을 예측하는 유효성이 크게 입증된 도구로 평가받아 국내·외에서 널리 사용되고 있다.^{1,2,20,25)}

한국판 음성활동 및 참여 프로파일(Korean Voice Activity and Participation Profile, K-VAPP)

Voice Activity and Participation Profile (VAPP)는 2001년 홍콩의 Ma와 Yiu²⁶⁾가 주축이 되어 국제기능건강분류(International Classification of Functioning, Disabilities, and Health)를 기반으로 만든 음성장애 관련 자기보고식 설문검사이다. 국내에서는 2016년 Lee 등²⁷⁾에 의해 K-VAPP가 번안되고 국내 음성장애 환자들을 대상으로 신뢰도와 타당도가 검증되었다. K-VAPP는 기존의 음성장애지수와 달리 음성장애로 인해 직장에서 겪는 어려움에 관한 문항이 더 추가되었으며, 일상 활동에서 받는 제한(activity limitation)과 실제로 활동에 참여하는 데서 오는 제약(participation restriction)을 분리하여 평가할 수 있도록 개발된 28문항으로 구성되어 있다.^{26,27)} 총점은 0점부터 280점까지이며 점수가 클수록 음성장애 환자가 경험하고 있는 일상활동에서 받는 제한이 큰 것으로 해석된다.^{26,27)} Lee 등²⁷⁾은 임상에서 음성장애 환자들의 선별을 위해 K-VAPP의 절단점을 총점(14.5점), 활동제한점수(4.5점), 참여제한점수(1.5점)로 각각 나누어 제시하였으며, 해당 절단점들은 우수한 민감도, 특이도를 나타내는 것으로 보고되었다. 또한 K-VAPP는 각 영역별 원점수들을 표준점수로 변환하여 산출할 수 있으며, 임상 현장에서의 활용도를 높이기 위해 각 점수별 백분위수를 제공하고 있다.²⁷⁾ K-VAPP는 직업 및 사회활동에서 음성 사용이 중요한 역할을 하는 환자들이 주관적으로 느끼는 목소리로 인한 불편을 보다 효과적으로 구분해낼 수 있는 도구로서 널리 활용되고 있다.^{26,27)}

한국어판 음성 파국화 지수(Korean Voice Catastrophization Index, K-VCI)

Voice Catastrophization Index (VCI)는 음성장애 환자들이 경험하고 있는 음성질환으로 비롯된 파국적 사고의 정도와 감정적 어려움을 측정하기 위해 2019년 Shoffel-Havakuk 등²⁸⁾이 개발한 평가도구이다. 국내에서는 2019년 Lee 등²⁹⁾에 의해 번안되고 신뢰도와 타당도가 검증되었다. 총 13개의 K-VCI 문항들은 음성장애 환자들이 경험하고 있는 반추(ruminization; 목소리 문제에 관해 지속해서 생각하는 정도), 과장(magnification; 목소리 관련 증세를 더 확대하여 심각하게 생각하는 정도), 무기력한(helplessness; 목소리 문제가 좋아지기 위해 할 수 있는 게 없다고 생각하는 정도) 사고의 정도를 평가하는 문항들로 구성되어 있고, 각 문항당 0-4점을 사용하여 평가한다.²⁹⁾ 총점은 0점부터 65점까지이며 점수가 높을수록 환자가 경험하고 있는 파국적 사고와 감정적 어려움이 큰 것으로 해석된다.^{28,29)} K-VCI의 절단점은 총점(6점), 반추(2점), 과장(3점), 그리고 무기력(3점)의 세 영역으로 나누어 Lee 등²⁹⁾에 의해 보고되었으며, 음성장애 환자가 겪고 있는 감정적인 어려움을 세부적으로 측정할 수 있는 도구로 확인되었다.

한국어판 음성관련 삶의 질 척도(Korean Voice-related Quality of Life, K-VQOL)

Voice-related Quality of Life (VQOL)은 미국의 Hogikyan과 Sethuraman³⁰⁾에 의해 음성장애 환자들이 경험하고 있는 삶의 질을 사회-심리 영역과 신체 영역으로 나누어 측정하기 위해 개발된 설문도구이다. 국내에서는 2007년 Kim 등²⁾에 의해 번안되고 국내 음성장애 환자들을 대상으로 신뢰도와 타당도가 입증되었다. 해당 척도는 목소리 문제가 생겼을 때에 삶의 질을 확인할 수 있는 10문항으로 구성되어 있으며, 환자들은 1-5점 척도를 사용하여 평가하게 된다.^{2,30)} K-VQOL은 점수 환산 공식을 사용하여 원점수를 0점부터 100점 사이로 표준화하고, 총점이 절단점인 91점보다 낮을수록 음성장애 환자가 경험하고 있는 음성관련 삶의 질이 나쁘다고 해석할 수 있다.^{2,9,30)} 특히 VQOL은 다양한 음성장애 임상집단에서 유효성이 입증되어 전 세계적으로 널리 사용되고 있는 도구이다.^{2,30)} K-VQOL은 간결한 문항으로 인해 짧은 시간 동안 음성장애 환자가 경험하고 있는 일상에서의 불편감을 쉽게 평가할 수 있으며, 음성질환의 치료 과정 동안 간편하게 반복 측정함으로써 환자가 보고하는 삶의 질의 변화 및 경과를 평가할 수 있는 도구로, 국내 임상현장에서도 그 유용성이 입증되었다.^{2,30)}

평가도구의 종합적 특성

본 연구에 포함된 14개 음성장애 환자용 자기보고식 평가도구의 전반적인 특성은 다음과 같다.

첫째, 국내 음성장애 관련 자기보고식 평가도구의 개발이 2016년도를 기점으로 활발히 이루어졌다는 점이다. 전체적으로 국내 음성장애 환자를 위한 자기보고식 평가도구들은 2007년부터 2021년까지 발표되었으며, 이 기간 동안 발표된 평가도구 중, 2016년 이후에 개발된 평가도구는 본 연구에 포함된 도구들의 절반 이상을 차지하였다. 이러한 동향은 The Organization for Economic Cooperation and Development가 환자 중심 보건의료를 독려하기 위해 2017년부터 시행하고 있는 Patient-Reported Indicators Survey 사업이 시작된 시점과 유사하며, 음성장애 관련 건강결과 측정이 기존의 자원과 의료이용 중심에서 환자가 직접 보고하는 '환자 경험 중심'으로 이동하고 있음을 보여준다.³¹⁾ 이러한 환자 중심의 건강결과 측정은 음성장애 치료가 환자의 개인적 수준에서 삶의 질, 기능, 신체, 감정적 영역에 미친 영향을 평가할 수 있게 해주며, 의료진 입장에서도 질환으로 인한 치료 결과가 환자가 느끼는 삶의 질에 미치는 영향을 추적할 수 있을 뿐만 아니라 다른 환자들의 치료경로, 치료계획과 비교할 수 있는 자료를 마련할 수 있게 해준다.

둘째, 음성장애 관련 자기보고식 평가도구의 다양화(diversification) 현상이다.^{5,6)} 2013년 이후로 음성질환이 일상 생활에 미치는 종합적인 영향을 평가하는 도구의 개발에서 나아가 수술 전후 목소리 변화, 목소리의 피로도, 목소리 문제로 인해 발생할 수 있는 사회적 제약, 감정적인 어려움과 같이 더 다양하고 세부적인 심리적 요인을 반영하는 도구들이 개발된 것이 확인되었다. 또한 2013년 이후로는 특정 질환군, 목소리 사용이 많은 직업군, 그리고 다양한 연령층을 위한 도구들로 대상자의 범위가 확장되어 설문도구들이 개발된 것을 알 수 있었다. 여기에는 갑상선 수술을 시행받은 환자들을 위해 개발한 TVSQ, 노년층을 위한 K-AVI와 AVI-KR, 그리고 엘리트 음성사용군인 가수들을 위해 개발된 K-SVHI와 같은 설문도구들이 포함되었다.^{10,12,13,24)} 이와 더불어 음성장애 아동을 위한 평가도구 중에서는 부모가 대신 작성하는 대리보고 형식에서 더 나아가 아동 스스로 경험하고 있는 목소리 관련 불편감을 측정할 수 있도록 개발된 K-CVHI-10과 같은 도구도 확인되었다.¹⁷⁾

셋째, 음성장애 관련 자기보고식 평가도구의 국가별 지역화(localization) 현상이다.¹⁾ 본 연구에 포함된 음성장애 관련 자기보고식 평가도구들은 국내 임상 현장, 문화, 사회적 배경을 반영하여 체계적으로 변안·개발되었음이 확인되었다. 특히 심리측정적 평가의 경우, 대상자의 언어, 문화와 같은 다

양한 요인들이 평가결과에 영향을 미친다는 것이 보고되어 있다.¹⁾ 따라서 국외 음성장애 환자 집단이 보이는 심리측정적 수치와 해석결과를 국내 환자 집단에 단순히 적용하는 것은 잘못된 해석을 유도할 우려가 있다. 그러므로 한국의 사회, 문화, 언어 등의 실정을 잘 반영한 평가도구들과 이러한 설문도구들을 국내 음성장애 환자 집단에 적용한 기준이 더욱 필요한 상황에서 음성장애 평가도구의 지역화는 의미 있는 현상으로 볼 수 있다. 이와 더불어 국내 실정에 맞게 개발되고 국내 음성장애 환자를 대상으로 심리측정적 속성이 검증되어 영문으로 발표된 TVSQ와 같은 설문도구가 국외에서도 사용되고 있는 현상은 상당히 고무적인 일이다.²⁴⁾ 이처럼 국내 실정을 반영하여 체계적인 검증절차를 걸쳐 개발된 자기보고식 평가도구들은 환자들이 인식하는 진료에 따른 기대치나 만족도를 이해하는데 중요한 정보를 제공해주며, 음성질환이 회복되는 과정에서 주요 증상에 대한 국내 환자들의 경험을 더 정확하게 측정하게 하는 데 기여할 수 있다.³¹⁾

임상 활용성

본 연구에 포함된 14개 평가도구의 임상 활용성을 분석한 결과 각 설문도구별로 차이가 있었으며, 종합적인 평가는 다음과 같다(Table 3).

첫째, 평가내용에 따른 설문도구 유형(type)을 비교하기 위해 개별 자기보고식 평가도구의 하위 평가항목, 문항별 내용들을 참고로 하여 각 설문도구를 4가지 유형(음성문제의 전반적인 불편을 측정하는 도구, 음성장애 증상을 평가하는 도구, 사회-감정적인 제약을 중심으로 측정하는 도구, 그리고 삶의 질 설문도구)으로 분류하였다. 14편의 문헌에서 가장 많이 변안·개발된 평가도구들은 음성장애 환자가 일상생활에서 느끼는 음성문제의 전반적인 불편감을 측정하는 도구들이었으며, 이들은 총 7개로 전체 평가도구의 절반(K-VHI, K-VHI-10, K-SVHI, K-AVI, AVI-KR, K-pVHI, K-CVHI-10)에 해당하였다.^{1,2,10,12,13,17)} 해당 평가도구들은 음성장애 환자가 느끼는 불편감을 기능, 신체, 정서적 측면으로 나누어 측정하도록 구성되어 있었다. 이와 더불어 음성장애와 관련된 신체적인 불편감을 집중하여 측정하는 도구들이 개발되었으며, 이러한 도구는 목소리의 피로도, 음성장애 환자들이 경험할 수 있는 목소리 관련 증상들, 갑상선 수술 전후 겪는 목소리 증상들을 측정할 수 있는 평가도구(n=4; K-VFI, K-VoiSS, K-VTDS, TVSQ)들로 구성되어 있었다.^{19,21,23,25)} 이 중 가장 문항 수가 많은 K-VoiSS는 목소리 관련 구체적인 증상뿐만 아니라 음성장애로 인해 겪을 수 있는 전반적인 어려움을 평가하는 문항들도 포함하고 있어 음성장애 환자의 경험을 가장 포괄적으로 측정할 수 있는 설문도구임을 알 수 있었다.^{18,19)} 또한 음

성장으로 인해 발생하는 사회적인 제약과 감정적 어려움을 중심으로 평가하는 도구들(n=2: K-VAPP, K-VCI)과 목소리 관련 삶의 질을 평가할 수 있는 자기보고 도구(n=1: K-VQOL)가 국내에서 개발된 것으로 확인되었다.^{2,27,29)}

둘째, 평가도구들의 적용 가능성(versatility)을 평가하기 위해 개별 평가도구들의 개발 시 포함된 음성장애 환자군의 특성과 해당 도구를 사용한 후속연구의 내용들을 바탕으로 각 설문도구를 3가지 범주(다양한 음성장애 환자군에 사용 가능한 도구, 특정 연령군 및 전문적 음성사용군에 사용 가능한 도구, 그리고 갑상선 수술을 받은 환자군에 사용 가능한 도구)로 분류하였다.

다양한 음성장애 환자군에 사용 가능한 도구로는 총 14개의 설문도구 중 8개의 평가도구들(57%, K-VHI, K-VHI-10, K-VFI, K-VoiSS, K-VTDS, K-VAPP, K-VCI, K-QOL)이 확인되었다.^{1,2,19,21,23,27,29)} 이와 더불어 특정 연령군, 직업적 음성사용군을 위한 평가도구들도 세부적으로 개발되어 사용되고 있었다. 구체적으로는 가수를 위한 음성장애 지수인 K-SVHI, 음성장애가 있는 국내 아동들을 위한 평가도구인 K-pVHI와 K-CVHI-10, 그리고 노년층을 위한 자기보고식 설문도구로는 원문이 AVI로 동일한 2개의 평가도구(K-AVI, AVI-KR)가 각각 다른 연구진들에 의해 번안·개발된 것으로 확인되었다.^{10,12,13,15,17)} 이와 더불어 갑상선 수술을 받은 환자들 경험할 수 있는 목소리 관련 증상을 구체적으로 평가할 수 있도록 개발된 TVSQ와 같은 도구가 보고되었으며, 이 외에도 K-VHI, K-VHI-10, 그리고 K-VoiSS가 갑상선 수술을 시행받은 환자군을 위한 설문도구로 사용 가능한 것으로 확인되었다.^{1,2,25)} 상대적으로 건강 상태, 일상생활 수행의 어려움, 경제 상황이 다를 수 있는 음성장애 환자군의 특성을 반영하는 설문도구의 개발은 음성장애의 정확한 진료와 치료를 가능케 해주며, 해당 집단의 특성을 반영한 자료를 확보하여 이후 비슷한 환자들을 평가하고 치료할 수 있는 근간이 되어 준다.

셋째, 각 음성장애 환자 자기보고식 평가도구들의 실용적인 특성(practicality)을 평가하기 위해 문항 수, 응답을 위해 사용된 척도 특성, 채점 방식, 그리고 절단점을 비교·분석한 결과를 살펴보면 다음과 같다. 문항 수의 경우 대부분의 평가도구들은 대략 20문항 정도로 구성되어 있다. 구체적으로는 평가도구들의 평균 문항 수는 21.2 (SD=10.7)이며, 적게는 8문항부터 많게는 44문항을 가진 설문도구가 개발되어 사용되고 있었다.^{21,23)} 그러나 각 설문도구를 작성하는 데 걸리는 평균시간은 모든 문헌에서 보고되지 않아 평가에서 제외되었다. 각 평가도구별 응답 방법을 비교한 결과, 대부분의 평가도구들은 음성장애 환자들이 답하기 쉽도록 Likert 척도를 사

용한 응답 방식을 사용하고 있었다. 그중 K-VAPP는 유일하게 11점 등간척도(equal-appearing interval scale)를 사용하였는데, 이는 시각 아날로그 척도(visual analog scale)의 특성을 반영한 것으로 보다 구체적으로 환자가 느끼는 음성장애로 인한 일상생활의 불편감을 평가하기 위한 연구자들의 노력으로 해석된다.^{26,27)} 채점방식의 경우 대부분의 평가도구들은 각 문항별 응답 값의 합으로 총점을 채점할 수 있으며, K-VAPP와 K-VQOL는 표준 점수를 산출하는 점수 계산 방법을 사용하고 있었다.^{2,27)} 절단점의 경우 6개(42.9%)의 평가도구만이 문헌을 통해 보고되었으며, 이 중 AVI-KR, K-CVHI-10, K-VAPP는 번안·개발 연구 당시 절단점이 함께 보고되었다.^{13,17,27)}

넷째, 평가도구들의 심리측정적 검증특성(psychometric features)을 살펴본 결과, 대부분의 도구는 음성장애로 진단 받은 환자군과 정상 목소리를 가진 이들을 대상으로 신뢰도와 타당도가 검증되었으며, 음성질환이 없는 직업적 음성사용인들과 일반인을 비교하여 검증한 도구도 보고되었다.¹⁹⁾ 내적일치도의 범위는 0.88부터 0.98까지로 높은 편이었으나 내적일치도가 보고되지 않은 도구도 확인되었다.¹⁰⁾ 각 평가도구들의 검사-재검사 신뢰도는 0.76부터 0.99까지로 양호한 편이었으며, 모든 설문도구들은 내용타당도(content validity), 구성타당도(construct validity), 공인타당도(concurrent validity)와 같은 통계방법을 사용하여 타당도 검증을 진행한 것을 알 수 있다. 그 외 안면타당도(face validity: TVSQ, K-VCI),^{25,29)} 요인타당도(factorial analysis: K-VHI, K-VHI-10)¹⁾를 추가로 분석하여 타당도를 검증한 설문도구들이 확인되어 대부분의 평가도구들이 체계적인 절차를 통해 신뢰도와 타당도가 검증된 것을 알 수 있다.

고 찰

본 연구는 음성장애 환자를 대상으로 개발된 총 14개의 자기보고식 평가도구의 일반적인 특성과 임상 활용성에 대해 분석하였다. 음성장애 환자가 직접 보고한 경험이 중요시되는 맥락 속에서 다양하고 세분화된 자기보고식 평가도구들이 번안·개발되어 사용되고 있음을 확인할 수 있었으며 대부분의 평가도구들은 체계적인 절차를 통해 신뢰도와 타당도가 검증된 설문 도구들임을 알 수 있었다.

그러나 본 연구의 분석 결과, 국내에서 개발된 음성장애 환자용 자기보고식 평가도구들의 정보 보고가 일부 미흡한 것을 확인할 수 있었다. 특히 모든 평가도구들의 연구 보고에 있어서 설문도구를 작성하는 데 걸린 평균시간을 보고하지 않은 점이 관찰되었다. 시간의 영향을 많이 받는(time-sen-

sitive) 임상 현장의 특성상 향후 설문도구를 개발하는 과정에서 해당 평가도구를 진료 시 사용하기에 용이하도록 설문을 작성하는 데 걸린 평균 시간을 보고할 필요성이 있다.⁶⁾ 이와 더불어 본 연구에 포함된 평가도구들 중 AVI-KR, K-CVHI-10, K-VAPP를 제외하고 대다수의 설문도구들이 개발 당시 절단점을 보고하지 않았다. 절단점은 음성장애 환자의 선별검사(screening test) 시 활용될 수 있고, 다면적인 음성 검사 결과와 함께 음성장애 환자를 종합적으로 진단하고 치료 예후를 파악하는 데 필요한 정보를 제공할 수 있으므로 향후 음성장애 관련 자기보고식 평가도구 개발 시 임상 현장에서 설문도구를 사용한 평가결과 해석이 용이하도록 절단점을 함께 보고하여 설문도구의 효용성을 뒷받침해야 할 것이다.²⁴⁾

절단점 보고와 더불어 국내에서 개발된 음성장애 관련 자기보고식 평가도구들의 추가검증을 통해 음성장애 환자의 치료 전후로 임상적으로 환자의 삶의 질과 증상개선에 의미 있는 변화가 일어났는지를 알 수 있는 ‘임상적으로 중요한 최소한의 변화값(minimal important difference, MID)’과 같은 정보도 제공되어야 할 것이다.⁶⁾ 특히 환자들의 주관적인 평가에 의존하는 자기보고식 평가도구의 특성상, 임상적으로도 환자에게 의미 있는 변화가 일어났는지에 대한 평가를 가능하게 해주는 MID를 보고하는 것은 음성장애 환자의 진료와 치료에 있어 각 설문도구들을 사용하는 근거를 마련해 줄 것이다.⁶⁾

마지막으로 하나의 평가도구를 사용하여 음성장애 환자의 삶의 질과 증상들을 포괄적으로 평가하기 어렵기 때문에, 임상가들은 각 설문도구가 어떤 음성장애 질환군의 증상 평가에 유용한지에 대한 정확한 지식을 갖고 임상 목표에 부합하는 적절한 설문도구를 채택하거나 각각 다른 측면을 평가할 수 있는 설문도구를 조합하여 사용해야 할 것이다. 만일 음성장애 환자들이 경험하고 있는 전반적인 목소리 관련 불편감과 신체적인 증상을 포괄적으로 평가해야 하는 경우 K-VHI 혹은 K-VoiSS를 사용할 수 있으며, 시간적인 부담을 최소화하기 위해서는 K-VHI-10이나 K-VQOL과 같이 문항 수가 적은 도구를 활용할 수 있을 것이다.^{1,2,19)} 평가하고자 하는 음성장애 환자가 직업적으로 목소리 사용량이 많은 음성 사용자에게 해당하는 경우, 음성질환으로 인해 발생할 수 있는 직업 및 사회활동의 제약을 구체적으로 측정할 수 있는 K-VAPP와 같은 설문도구를 K-VHI, K-VHI-10, K-VQOL과 같이 사용함으로써 환자들의 경험하고 있는 어려움을 다면적으로 평가할 수 있을 것으로 사료된다.^{1,2,27)}

결론

본 연구는 국내에서 번안·개발된 총 14개의 음성장애 관련 자기보고식 평가도구들의 특성과 실용성을 분석한 결과, 환자 중심의 건강보고가 중요시되는 흐름 속에서 다양하고 세분화된 음성장애 환자를 위한 자기보고식 설문도구들이 국내 실정에 맞게 번안·개발되어 사용되고 있음을 확인하였다. 아울러 본 연구는 향후 평가도구 개발 시 보완되어야 할 내용들뿐만 아니라 각 평가도구가 어떤 환자군과 증상 평가에 유용한지에 대한 분석정보를 기반으로 진료 목적에 적합한 설문 도구를 사용할 것을 제안하고 있다.

Acknowledgments

This research was supported by Post-Doctoral Fellowship Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2021R1A6A01087023).

Author Contribution

Conceptualization: Jeong Min Lee, Han Su Kim. Data curation: Jeong Min Lee. Formal analysis: Jeong Min Lee. Funding acquisition: all authors. Investigation: all authors. Methodology: Jeong Min Lee, Han Su Kim. Project administration: all authors. Supervision: Soo Yeon Jung, Han Su Kim. Validation: Han Su Kim. Visualization: Jeong Min Lee, Han Su Kim. Writing—original draft: Jeong Min Lee. Writing—review & editing: Soo Yeon Jung, Han Su Kim.

ORCIDiDs

Han Su Kim <https://orcid.org/0000-0003-2239-0225>
 Jeong Min Lee <https://orcid.org/0000-0003-0279-0731>
 Soo Yeon Jung <https://orcid.org/0000-0001-7497-3057>

REFERENCES

- 1) Yun YS. Korean-voice handicap index (K-VHI): Validation of an original version and development of new version [dissertation]. Seoul: Yonsei Univ.;2007.
- 2) Kim JO, Lim SE, Park SY, Choi SH, Choi JN, Choi HS. Validity and reliability of Korean-version of voice handicap index and voice-related quality of life. *Speech Science* 2007;14(3):111-25.
- 3) Jacobson BH, Johnson A, Grywalski C, Silbergleit A, Jacobson G, Benninger MS, et al. The voice handicap index (VHI) development and validation. *Am J Speech Lang Pathol* 1997;6(3):66-70.
- 4) Arksey H, O'Malley L. Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol* 2005;8(1):19-32.
- 5) Pyo HY. A current trend of self-reporting evaluation tools for voice disorders. *Audiol Speech Res* 2019;15(3):160-7.
- 6) Francis DO, Daniero JJ, Hovis KL, Sathe N, Jacobson B, Penson DF, et al. Voice-related patient-reported outcome measures: A systematic review of instrument development and validation. *J Speech Lang Hear Res* 2017;60(1):62-88.
- 7) Lee YW, Kim GH. Discriminative and predictive ability for screening the Korean dysphonic patients using self-reported questionnaires. *Clin Arch Commun Disord* 2020;5(2):85-95.
- 8) Rosen CA, Lee AS, Osborne J, Zullo T, Murry T. Development and validation of the voice handicap index-10. *Laryngoscope* 2004; 114(9):1549-56.
- 9) Cohen SM, Jacobson BH, Garrett CG, Noordzij JP, Stewart MG,

Attia A, et al. Creation and validation of the singing voice handicap index. *Ann Otol Rhinol Laryngol* 2007;116(6):402-6.

10) Lee AR, Sim HS. The Korean version of the singing voice handicap index. *Commun Sci Disord* 2013;18(2):194-202.

11) Etter NM, Hapner ER, Barkmeier-Kraemer JM, Gartner-Schmidt JL, Dressler EV, Stemple JC. Aging voice index (AVI): Reliability and validity of a voice quality of life scale for older adults. *J Voice* 2019;33(5):807.e7-12.

12) Bae IH, Sung ES, Lee JC. Validity and reliability of Korean version of the aging voice index (KAVI). *J Korean Soc Laryngol Phoniatr Logop* 2019;30(1):21-7.

13) Pyo HY, Lim JY, Lim SE, Lee SJ. Comparison of voice-related quality of life for the elderly with and without voice disorders according to genders by aging voice index-Korean version. *Commun Sci Disord* 2021;26(4):933-43.

14) Zur KB, Cotton S, Kelchner L, Baker S, Weinrich B, Lee L. Pediatric voice handicap index (pVHI): A new tool for evaluating pediatric dysphonia. *Int J Pediatr Otorhinolaryngol* 2007;71(1):77-82.

15) Park SS, Kwon TK, Choi SH, Lee WY, Hong YH, Jeong NG, et al. Reliability and validity of the Korean version of pediatric voice handicap index: In school age children. *Int J Pediatr Otorhinolaryngol* 2013;77(1):107-12.

16) Ricci-Maccarini A, De Maio V, Murry T, Schindler A. Development and validation of the children's voice handicap index-10 (CVHI-10). *J Voice* 2013;27(2):258.e23-8.

17) Ra SH. Reliability and validity of the Korean version of children's voice handicap index-10 (CVHI-10) [dissertation]. Daegu: Catholic University of Daegu;2015.

18) Deary IJ, Wilson JA, Carding PN, MacKenzie K. VoiSS: A patient-derived voice symptom scale. *J Psychosom Res* 2003;54(5):483-9.

19) Son HY, Lee CY, Kim KA, Kim S, Jeong HS, Kim JP. The Korean version of the voice symptom scale for patients with thyroid operation, and its use in a validation and reliability study. *J Voice* 2018;32(3):367-73.

20) Nanjundeswaran C, Jacobson BH, Gartner-Schmidt J, Verdolini Abbott K. Vocal fatigue index (VFI): Development and validation. *J Voice* 2015;29(4):433-40.

21) Kim MJ. Relationship between vocal fatigue index and voice handicap index for professional voice users [dissertation]. Seoul: Ewha Womans Univ.;2016.

22) Mathieson L. Vocal tract discomfort in hyperfunctional dysphonia. *J Voice* 1993;2:40-8.

23) Lee HN, Yoo JY, Han JH, Park YS, Jung DY, Park JH. Transcultural adaptation and validation of the Korean version of the vocal tract discomfort scale. *J Voice* 2022;36(1):143.e15-23.

24) Nam IC, Bae JS, Shim MR, Hwang YS, Kim MS, Sun DI. The importance of preoperative laryngeal examination before thyroidectomy and the usefulness of a voice questionnaire in screening. *World J Surg* 2012;36(2):303-9.

25) Hwang YS, Shim MR, Kim GJ, Lee DH, Nam IC, Park JO, et al. Development and validation of the thyroidectomy-related voice and symptom questionnaire (TVSQ). *J Voice* 2022;36(1):145.e15-22.

26) Ma EP, Yiu EM. Voice activity and participation profile: Assessing the impact of voice disorders on daily activities. *J Speech Lang Hear Res* 2001;44(3):511-24.

27) Lee SJ, Choi HS, Kim HH, Byeon HK, Lim SE, Yang MK. Korean version of the voice activity and participation profile (K-VAPP): A validation study. *Commun Sci Disord* 2016;21(4):695-708.

28) Shoffel-Havakuk H, Chau S, Hapner ER, Pethan M, Johns MM 3rd. Development and validation of the voice catastrophization index. *J Voice* 2019;33(2):232-8.

29) Lee YW, Kim GH, Bae IH, Park HJ, Kwon SB. Korean voice catastrophization index (K-VC1): Validation of the voice catastrophization index for Koreans. *J Voice* 2019;35(1):160.e7-13.

30) Hogikyan ND, Sethuraman G. Validation of an instrument to measure voice-related quality of life (V-RQOL). *J Voice* 1999; 13(4):557-69.

31) Choi JS, Park YS, Kim J, Park CS. International trends on patient-reported outcome measures for improving care quality and its implication for South Korea: Focus on OECD PaRIS. *Qual Improv Health Care* 2019;25(1):11-28.

정답 및 해설

답 ④

해설 Immunoglobulin G4 연관 질환(immunoglobulin G4 related disease, IgG4-RD)은 혈청 IgG4 상승, IgG4 양성 형질세포 및 림프구 침윤, 섬유화를 특징으로 하는 질환이다. IgG4 연관 질환은 비교적 최근에 알려진 질환으로 섬유염증성 병변과 연관된 면역 매개 질환이며, 인체의 모든 장기에 침범될 수 있으며, 이비인후과 영역에서는 악하선, 이하선, 설하선 등 주요 침샘이나 눈물샘, 갑상선 등에 발생할 수 있고 드물게 비부비동을 침범할 수 있다. 따라서 진단이 될 경우 각 장기의 침범 여부를 확인해야 하며, 치료는 코르티코스테로이드(corticosteroid treatment)가 기본이 된다.

참고 문헌: Han YS, Cha JS, Jung MJ, Kim JY. A case of IgG4-related disease with sino-nasal involvement. *Korean J Otorhinolaryngol-Head Neck Surg*. In press 2022.

Flint P, Haughey B, Lund V, Robbins K, Thomas JR, Lesperance M, et al. *Cummings otolaryngology head and neck surgery*. 7th ed. Elsevier;2020. p.2309.