



Acoustic and Auditory-Perceptual Evaluation as Predictor of Voice Recovery after Laryngeal Microsurgery in Patients with Vocal Polyp

Geun-Hyo Kim¹, Yeon-Yoo Lee¹, Byung-Joo Lee¹, and Soon-Bok Kwon²

¹Department of Otorhinolaryngology-Head and Neck Surgery, Pusan National University Hospital, Busan; and

²Department of Humanities, Language and Information, Pusan National University, Busan, Korea

성대용종 환자의 후두미세수술 후 음성회복의 예측인자로서의 음향학적 및 청지각적 평가

김근효¹ · 이연우¹ · 이병주¹ · 권순복²

부산대학교병원 이비인후과,¹ 부산대학교 언어정보학교실²

Background and Objectives The purpose of this study was to investigate acoustic (acoustic voice quality index, AVQI) and auditory-perceptual evaluation (grade, rough, breathy, asthenic, strained; GRBAS and consensus auditory-perceptual evaluation of voice; CAPE-V) as a predictor of voice recovery after laryngeal microsurgery (LMS) in patients with vocal polyp.

Subjects and Methods A total of 68 patients with vocal polyp participated in this study. Voice samples were analyzed for AVQI by Praat and auditory-perceptual ratings were performed by three speech language pathologists. Voice handicap index-10 (VHI-10) was evaluated by patients themselves.

Results Decreased AVQI, VHI-10, overall severity (OS), and increased smoothed cepstral peak prominence (CPPS) values were measured and statistically significant changes were noted after LMS. The ratio of Grade 0 and Grade 1 was increased. AVQI was correlated with CPPS, Grade, and OS, but not with VHI-10. The voice recovery of pedunculated polyp appeared in all vocal polyp sizes after LMS, but statistically significant differences were found only in small and medium sizes. After LMS of sessile polyps, AVQI, VHI-10, and OS decreased whereas CPPS increased; however, statistically significant difference was confirmed only in VHI-10 and OS. The receiver operating characteristic curve analysis showed the following results: AVQI [cutoff=5.5, sensitivity=61.8%, specificity=76.5%, area under the curve (AUC)=0.712], CPPS (cutoff=5.0, sensitivity=89.7%, specificity=48.5%, AUC=0.743), VHI-10 (cutoff=13.0, sensitivity=77.9%, specificity=89.7%, AUC=0.893), Grade (cutoff=2.0, sensitivity=70.6%, specificity=69.1%, AUC=0.728), and OS (cutoff=51.0, sensitivity=86.8%, specificity=66.2%, AUC=0.833).

Conclusion Acoustic and auditory-perceptual variables showed significant positive changes and predicted recovery of voice. In this study, we believe that the acoustic and auditory-perceptual evaluations.

Korean J Otorhinolaryngol-Head Neck Surg 2018;61(7):361-9

Key Words Acoustic voice quality index · Consensus auditory-perceptual evaluation of voice · Grade, rough, breathy, asthenic, strained · Praat · Vocal polyp.

Received January 4, 2018

Revised January 29, 2018

Accepted February 9, 2018

Address for correspondence

Soon-Bok Kwon, PhD

Department of Humanities,

Language and Information,

Pusan National University,

2 Busandaehak-ro 63beon-gil,

Geumjeong-gu, Busan 46241, Korea

Tel +82-51-510-2003

Fax +82-51-582-0572

E-mail sbkwon@pusan.ac.kr

서 론

성대용종은 성인 후두질환 중 흔하게 관찰되며, 주로 출혈성 용종은 남자에게 많이 발생한다. 순간적으로 강한 발성, 과도한 성대 마찰을 유발하는 발성 등이 주된 원인이며, 이외에도 잦은 흡연과 음주생활, 항응고제 복용, 만성 위산역류 등으로 인한 후두자극 등이 영향을 줄 수 있다. 병리조직학적으로는 성대결절과 유의한 차이는 없으며, 육안으로 확인하였을 때 점액성(muroid type)과 혈관종성(angiomatus type)으로 나눌 수 있다. 과도한 발성으로 인하여 미세혈관 파열이 나타나고 점막하 공간에 출혈을 일으키게 된다. 이러한 출혈이 국소적으로 제한되어 커지면 유경형 성대용종(pedunculated polyp)이 발생하며, 넓은 범위에 발생하는 경우는 범발성 성대용종(sessile polyp)이 된다. 대부분 편측성으로 발생하며 반대 측 성대 부위에 반응성 결절(reactive nodule)이 생기는 경우도 많다. 치료방법으로는 보존적 치료 방법으로 음성의 오남용을 줄이고, 금주, 금연 등 생활습관 교정, 음성위생 등과 함께 단기간의 음성치료를 통해서 호전되는 경우가 많다. 만성적이고 큰 용종 또는 3개월 이상의 음성치료에도 성대용종의 크기에 변화가 없다면 수술적 치료를 고려해야 한다. 후두미세수술 시에는 수술 후의 음성회복을 위해서 성대 점막 및 점막 하부의 정상 조직을 보존하는 것이 필수적이다.^{1,2)}

후두미세수술 후 음성회복을 확인하기 위해서는 음성평가를 시행해야 한다.^{3,4)} 병리적인 음성에 대한 음향학적 음성평가는 비침습적인 검사이며 성대질환의 치료 전후 회복 정도를 정량화 할 수 있어 필수적으로 시행되고 있다. 일반적으로 음향학적 음성평가를 위해서 모음연장발성 및 문장 읽기(산책문단, 가을문단)의 과업을 수행한다. 모음연장발성은 음성분석 프로그램인 multi-dimensional voice program(MDVP; model 5105, KayPentax, Montvale, NJ, USA)을 이용하여 기본주파수, 주파수변동률, 강도변동률, 소음 대 배음비 등을 측정한다. 또한, 문장 읽기는 Computerized speech lab(CSL; model 4500, KayPentax)을 이용하여 발화주파수(speaking fundamental frequency)를 측정한다. 매우 심한 발성장애를 가진 환자들은 analysis of dysphonia in speech and voice (ADSV, model 5109, KayPentax)와 SpeechTool 등으로 캡스트럼 분석을 실시하여 cepstral peak prominence(CPP) 혹은 cepstral peak prominence smoothed(CPPS)를 측정한다.

기존의 음성분석은 모음연장발성과 문장 읽기 과업을 각각 분석하여 발성장애의 정도를 확인하였다. 두 과업 간에도 발성장애 정도가 다른 경우가 많아서 종합적인 음성평가의 결과 해석에 어려움을 겪기도 한다. 최근 두 과업을 함께 분석하여 발성장애의 정도를 알 수 있는 acoustic voice quality

index(AVQI)라는 측정방법이 소개되었다.^{5,6)} 섭동 분석과 캡스트럼 분석이 모두 가능한 Praat(Institute of Phonetic Sciences, University of Amsterdam, Amsterdam, the Netherlands)을 통해서 측정된다. 섭동 분석과 캡스트럼 분석에서 측정되는 변수들에 가중치를 준 회귀식을 통해서 AVQI가 계산된다. 가장 중요한 특징은 모음과 문장 읽기를 통합해서 분석하며, 쉼과 무성음 등을 자동으로 제거하여 음성의 질을 정량화한다는 것이다. 다양한 측정변수들을 Praat 스크립트를 통해서 대량의 음성샘플을 자동으로 분석할 수 있게 되었다. AVQI의 회귀식은 신뢰성 높은 정량화를 위해서 지속적으로 보완되고 있다.

다양한 언어권에서도 AVQI의 유효성이 보고되었으며, 임상현장에서 이루어지고 있는 청지각적 평가(grade, rough, breathy, asthenic, strained; GRBAS와 consensus auditory-perceptual evaluation of voice; CAPE-V)와도 높은 상관관계가 확인되었다.⁵⁻¹¹⁾ 두 과업의 발성 특징들이 모두 반영되는 AVQI를 통하여 후두미세수술 전후의 음성평가를 실시하였다. 본 연구에서는 AVQI와 청지각적 평가를 이용하여 성대용종 환자의 후두미세수술 전후의 음성을 비교하고 음성의 회복을 확인하고자 한다.

대상 및 방법

대 상

2016년 6월부터 2017년 12월까지 본원에 내원한 환자 중에서 후두내시경, 후두스트로보스코피 검사상 성대용종으로 진단받고 후두미세수술을 받은 69명 대상자 중 신호 대 잡음비(signal to noise ratio, SNR) 평가를 통해서 30 dB 이하인 1명(27.9 dB)의 환자를 제외한 68명을 대상으로 하였다. 음성분석의 높은 신뢰도를 확보하기 위해서는 신호 대 잡음비(SNR)가 최소한 30 dB 이상은 확보되어야 한다.^{12,13)} 대상자들에 대한 정보를 Table 1에 제시하였다. 성대용종의 크기는

Table 1. Primary laryngostroboscopic diagnoses of subjects included in the study

	Pedunculated type (n=39)	Sessile type (n=29)
Sex	14:25	18:11
Age	48.7±12.9	47.5±15.0
Lesion position		
Both	5	2
Right	18	18
Left	16	9
Size	4 (large), 25 (medium), 10 (small)	Broad

Fig. 1. Classification of vocal polyp. pedunculated vocal polyp (A). sessile vocal polyp (B).

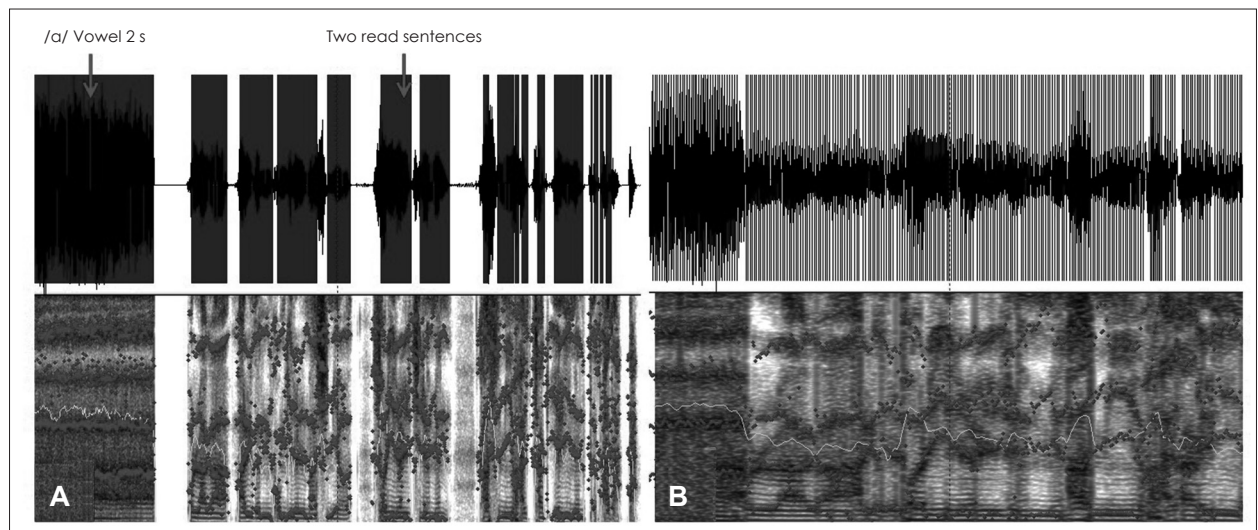
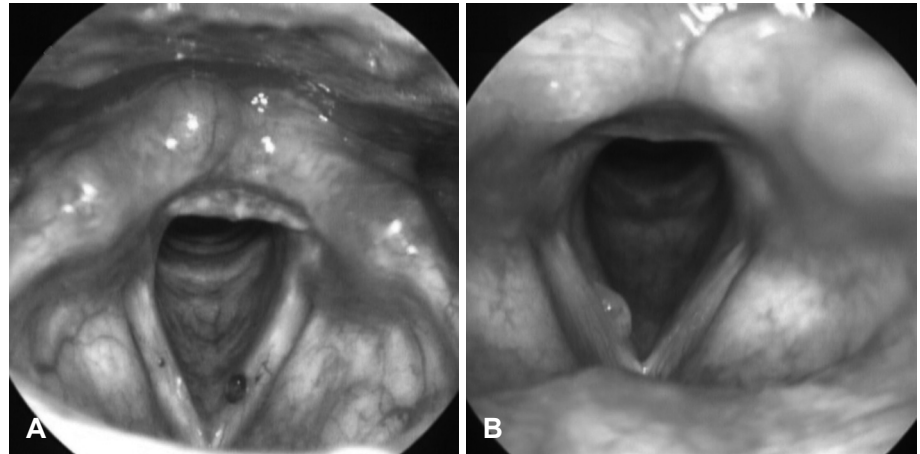


Fig. 2. Oscillographic display (y-axis: sound pressure in Pascal) and narrowband-spectrogram (y-axis: frequency in Hz, window length=0.05 s) of a concatenated voice sample. As used for the perceptual evaluations (A). As used for the acoustic evaluations, Praat script of Maryn, et al.⁶⁾ for automated detection, segmentation, and concatenation of these voiced segments was applied (B).

Jeong 등¹⁴⁾의 논문에서 제시한 기준으로 분류하였다(large: 전체 성대길이의 1/3 이상 보다 큼, medium: 전체 성대길이의 1/3보다 작음, small: 아주 작음(pinpoint). 후두미세수술 후 성대용종의 분류에 따른 음성변화를 관찰하기 위해서 유경형, 범발성 성대용종으로 분류하였다(Fig. 1).¹⁵⁾ 본 연구는 이미 획득된 음성 샘플을 후향적으로 분석한 연구로서 부산대학교병원 연구윤리위원회(Institutional Review Board)의 심의를 면제받았다(IRB No. H-1801-012-063).

음성녹음 및 편집

음성녹음은 이비인후과 음성검사실 내 방음 부스에서 음성 분석기기인 CSL 4500(Kaypentax)을 사용하여 두 가지 과업, 모음연장발성과 '산책' 문단 중 2문장 읽기를 실시하였다. 모음연장발성은 3회 반복하여 안정적으로 발성한 샘플의 시작과 끝부분을 제외한 안정구간 1.5초를 선정하고, 문장 읽기는

2개의 문장을 편안한 속도, 강도로 읽도록 지시하였고, 샘플링 레이트 44.1 kHz, 양자화 16 bit, WAV 형식으로 녹음하였다.

AVQI는 모음연장발성과 문장 읽기 샘플을 하나의 음성파일로 합성하여 측정하게 된다. Praat 스크립트를 통해서 두 과업의 음성샘플을 연결하고, 쉼, 무성음 구간을 제거하게 된다(Fig. 2). 본 연구에서는 선행연구에서 제시한 스크립트를 본 연구의 목적에 맞게 수정, 보완하여 적용하였다.⁶⁾

음향학적 평가 및 음성설문지

성대용종의 후두미세수술 전후의 음성에 대한 음향학적 평가를 위해서 AVQI를 측정하였다. AVQI는 캡스트럴 피크 현저성(CPPS), 배음 대 소음비(harmonics to noise ratio, HNR), shimmer local(SL), shimmer local dB(SLdB), 장기 평균 스펙트럼의 일반 기울기(general slope of the long-term average spectrum, Slope), 장기 평균 스펙트럼을 통한 회귀선 기울기

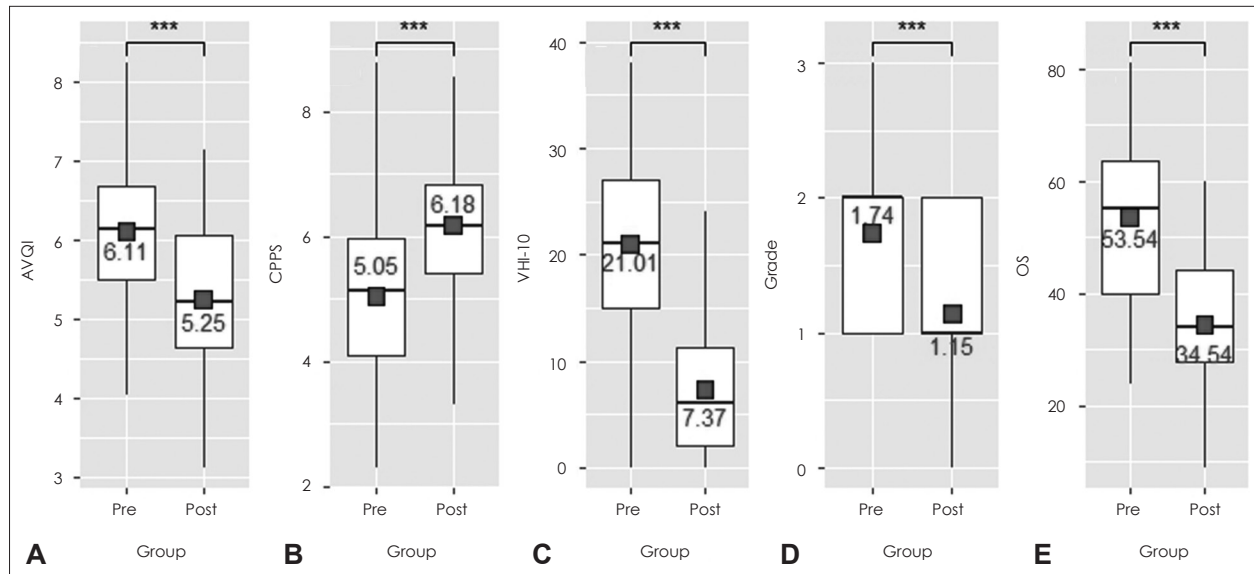


Fig. 3. Boxplots for pre/post operative evaluation of voice. AVQI (A), CPPS (B), VHI-10 (C), Grade (D), and OS (E). AVQI: acoustic voice quality index, CPPS: cepstral peak prominence smoothed, OS: overall severity, VHI-10: voice handicap index-10.

(tilt of the regression line through the long-term average spectrum, Tilt)를 사용하여 측정된다. AVQI 측정식(version 0.202)은 다음과 같다.^{9-11,16)}

$$9.072 - (0.245 \times \text{CPPS}) - (0.161 \times \text{HNR}) - (0.470 \times \text{SL}) + (6.158 \times \text{SLdB}) - (0.071 \times \text{Slope}) + (0.170 \times \text{Tilt})$$

또한, 환자 스스로 본인의 음성을 평가하기 위해서 voice handicap index-10(VHI-10)을 후두미세수술 술 전, 술 후에 작성하였다.

청지각적 평가

음성평가 및 치료의 경력이 7년 이상인 언어재활사 3명이 음성샘플을 듣고 평가하였다. 대상자들에 대한 사전 정보없이 청지각적 평가를 수행하였으며 음성샘플을 무작위로 제시되었다. GRBAS 척도에서 전체적인 상태를 나타내는 Grade(G)를 4점 척도(0: normal, 1: mild, 2: moderate, 3: severe), CAPE-V에서는 Grade에 상응하는 overall severity(OS)를 Visual Analysis Scale로 평가하였다. 이에 따른 평가자 간 신뢰도는 87%였으며, 평가자 내 신뢰도 분석을 위하여 전체 대상자 중 20%에 해당하는 14명의 음성샘플을 무작위로 선택하여 첫 평가 후 2주 내에 동일한 평가자가 재평가를 실시하였다. 이에 따른 평가자 내 신뢰도는 91%로 나타났다.

자료처리

본 연구의 자료는 R, version 3.4.1(The R Foundation for

Statistical Computing, Vienna, Austria)과 RStudio 1.0.143(RStudio Inc., Boston, MA, USA)을 이용하여 분석되었다. 정규성 검정을 만족하여 후두미세수술 전, 후의 음성결과 차이는 대응표본 t-검정(paired t-test)을 통해 유의성을 검증하였다. 유크형(large) 상대용종은 정규성 검정을 만족하지 못하여 월콕슨 부호 순위검정(Wilcoxon signed rank test)을 사용하여 유의성을 검증하였다.

AVQI, CPPS, G, OS, VHI-10 간의 상관분석은 피어슨 상관분석(Pearson correlation coefficient)을 통해서 검증하였으며 유의수준은 95%였다. 이때의 상관계수(r)은 -1에서 1 사이의 값을 가지며 양수는 양의 상관관계, 음수는 음의 상관관계를 나타낸다.

후두미세수술 전후의 음성평가 결과에 대한 최적의 절단값(cutoff) 및 정확도(area under the area, AUC)를 찾기 위해서 receiver operating characteristic(ROC) 곡선 분석을 실시하였다. Epi, pROC 패키지를 사용하여 ROC 분석 및 그림 작업을 실시하였다.

결 과

후두미세수술 전후 음성평가 비교

후두미세수술 전후 음성 비교를 위해서 AVQI, CPPS, G, OS, VHI-10을 평가하였다. 각 변수들의 후두미세수술 전후 비교 결과는 Fig. 3과 Table 2에 제시하였다. AVQI의 값을 비교한 결과, 술 전 6.1 ± 1.0 , 술 후 5.3 ± 1.0 으로 측정되어 술 전에 비해 술 후 AVQI값이 통계적으로 유의하게 낮았다($p <$

0.01). CPPS의 값을 비교한 결과, 술 전 5.0 ± 1.3 , 술 후 6.2 ± 1.1 로 측정되어 술 전에 비해 술 후 CPPS값이 통계적으로 유의하게 높았다($p < 0.01$). VHI-10의 값을 비교한 결과, 술 전 21.0 ± 8.5 , 술 후 7.4 ± 6.2 로 측정되어 술 전에 비해 술 후 VHI-10 값이 통계적으로 유의하게 낮았다($p < 0.01$). Grade의 값을 비교한 결과, 술 후의 Grade는 술 전에 비해 좋은 음성치의 비율이 높게 변화하였다($p < 0.01$). 술 전에 G2, G3의 비

율이 높았지만 술 후에 G1, G2의 비율이 증가하였다. OS의 값을 비교한 결과, 술 전 53.5 ± 13.8 , 술 후 34.5 ± 13.7 로 측정되어 술 전에 비해 술 후 OS 값이 통계적으로 유의하게 낮았다($p < 0.01$).

측정 변수간 상관관계

술 전 변수들의 상관관계를 분석한 결과, AVQI는 OS($r = 0.884, p < 0.01$), CPPS($r = -0.799, p < 0.01$), G($r = 0.749, p < 0.01$)와 높은 상관관계를 보였으며 VHI-10($r = 0.179, p = 0.142$)과는 낮은 상관관계를 보였다. CPPS는 OS($r = 0.746, p < 0.01$), G($r = 0.645, p < 0.01$), VHI-10($r = -0.294, p < 0.05$)과 상관관계를 보였다. Grade는 OS($r = 0.836, p < 0.01$), VHI-10($r = -0.347, p < 0.01$)과 상관관계를, OS는 VHI-10($r = 0.345, p < 0.01$)과 상관관계를 보였다.

술 후 변수들의 상관관계를 분석한 결과, AVQI는 CPPS($r = -0.868, p < 0.01$), OS($r = 0.800, p < 0.01$), G($r = 0.703, p < 0.01$)와 높은 상관관계를 보였으며 VHI-10($r = 0.187, p = 0.126$)과는 낮은 상관관계를 보였다. CPPS는 OS($r = 0.809, p < 0.01$), G($r = 0.721, p < 0.01$), VHI-10($r = -0.242, p < 0.05$)과 상관관계를 보였다. Grade는 OS($r = 0.845, p < 0.01$), VHI-10($r = -0.188, p = 0.125$)과 상관관계를, OS는 VHI-10($r = 0.305, p < 0.05$)과

Table 2. Results of voice evaluation pre and post LMS in patients with vocal polyps

Parameter	Pre	Post	t	p-value*
AVQI	6.1 ± 1.0	5.3 ± 1.0	7.988	<0.001
CPPS	5.0 ± 1.3	6.2 ± 1.1	-1.411	<0.001
VHI-10	21.0 ± 8.5	7.4 ± 6.2	14.493	<0.001
Grade (%)			8.746	<0.001
0	0 (0.0)	10 (14.7)		
1	21 (30.9)	38 (55.9)		
2	44 (64.7)	20 (29.4)		
3	3 (4.4)	0 (0.0)		
OS	53.5 ± 13.8	34.5 ± 13.7	14.977	<0.001

*significant at 0.01 level (two-tailed). AVQI: acoustic voice quality index, CPPS: cepstral peak prominence smoothed, LMS: laryngeal microsurgery, OS: overall severity, Pre: pre laryngeal microsurgery, Post: post laryngeal microsurgery, VHI-10: voice handicap index-10

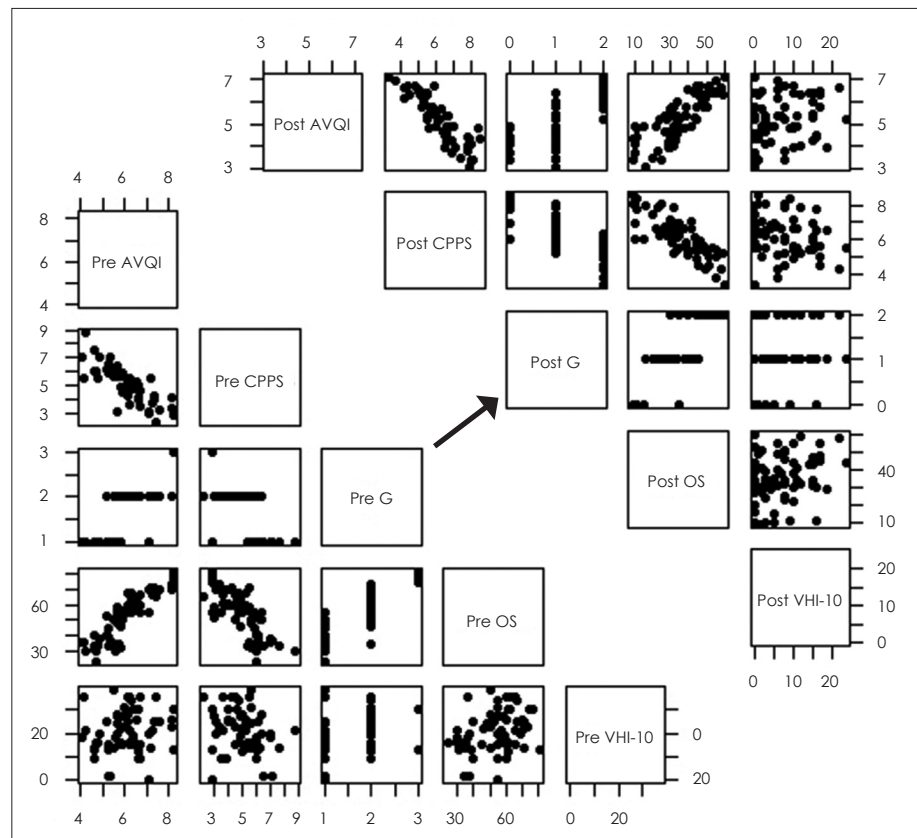


Fig. 4. Correlation plots among measured parameters in pre/post operative evaluation. AVQI: acoustic voice quality index, CPPS: cepstral peak prominence smoothed, G: grade, OS: overall severity, VHI-10: voice handicap index-10.

Table 3. Results of voice evaluation post LMS according to size of pedunculated polyps

	Small			Medium			Large		
	Pre (n=10)	Post (n=10)	p-value	Pre (n=25)	Post (n=25)	p-value	Pre (n=4)	Post (n=4)	p-value
AVQI	6.2±1.3	5.2±0.8	0.044*	6.3±1.1	5.5±1.0	0.006†	6.1±0.6	5.3±1.5	0.375
CPPS	4.9±1.4	6.0±1.0	0.072	4.7±1.1	5.9±1.1	0.001†	4.9±0.7	5.9±1.1	0.194
VHI-10	18.6±7.5	7.9±7.2	0.000†	24.4±7.2	8.6±6.0	0.000†	20.0±8.7	7.5±2.6	0.000†
Grade (%)			—			0.004†			—
0	0 (0.0)	0 (0.0)		0 (0.0)	3 (12.0)		0 (0.0)	0 (0.0)	
1	3 (30.0)	8 (80.0)		4 (16.0)	13 (52.0)		1 (25.0)	2 (50.0)	
2	7 (7.0)	2 (20.0)		19 (76.0)	9 (36.0)		3 (75.0)	2 (50.0)	
3	0 (0.0)	0 (0.0)		2 (8.0)	0 (0.0)		0 (0.0)	0 (0.0)	
OS	55.1±12.8	36.2±9.6	0.001†	56.9±13.0	37.3±15.1	0.000†	57.8±9.0	35.8±7.4	0.009†

*significant at 0.05 level (two-tailed), †significant at 0.01 level (two-tailed). AVQI: acoustic voice quality index, CPPS: cepstral peak prominence smoothed, LMS: laryngeal microsurgery, OS: overall severity, Pre: pre laryngeal microsurgery, Post: post laryngeal microsurgery, VHI-10: voice handicap index-10

Table 4. Results of voice evaluation in sessile polyps post LMS

	Sessile		
	Pre (n=29)	Post (n=29)	p-value
AVQI	6.0±1.0	5.5±1.1	0.069
CPPS	5.2±1.3	5.7±1.1	0.094
VHI-10	21.3±7.5	7.8±6.5	0.000*
Grade (%)			—
0	0 (0.0)	5 (17.0)	
1	9 (31.0)	10 (34.0)	
2	19 (65.5)	14 (48.3)	
3	1 (3.4)	0 (0.0)	
OS	53.2±13.8	37.5±14.9	0.000*

*significant at 0.01 level (two-tailed). AVQI: acoustic voice quality index, CPPS: cepstral peak prominence smoothed, LMS: laryngeal microsurgery, OS: overall severity, Pre: pre laryngeal microsurgery, Post: post laryngeal microsurgery, VHI-10: voice handicap index-10

상관관계를 보였다(Fig. 4).

유경형 성대용종(Pedunculated type)의 병변 크기에 따른 음성평가 비교

유경형 성대용종의 병변 크기에 따른 음성평가를 비교한 결과를 Table 3에 제시하였다. 병변 크기에 따른 후두미세수술 전후 비교 결과, 모든 성대용종에서 감소된 AVQI, VHI-10, OS 값, 증가된 CPPS 값이 관찰되었다. Grade의 경우 모든 성대용종에서 G2, G3 값의 비율이 감소하고 G1, G2의 비율이 증가하였다.

범발성 성대용종(Sessile type)의 수술 전후의 음성평가 비교

범발성 성대용종의 수술 전후 음성평가의 결과를 Table 4에 제시하였다. 후두미세수술 전후 비교 결과, 술 후의 음성은 술 전에 비해 감소된 AVQI, VHI-10, OS 값, 증가된 CPPS

값이 관찰되었고, VHI-10, OS에서만 통계적으로 유의한 차이가 확인되었다. Grade의 경우 술 후 G2, G3의 비율이 감소하고 G0의 비율이 증가하였다.

후두미세수술 전후 음성에 대한 ROC 곡선 분석

후두미세수술 전후의 음성에 대한 최적의 절단값과 정확도를 구하기 위하여 ROC 곡선 분석을 실시한 결과는 Fig. 5에 제시하였다. ROC 곡선 분석 결과, AVQI(cutoff=5.5, sensitivity=61.8%, specificity=76.5%, AUC=0.712), CPPS(cutoff=5.0, sensitivity=89.7%, specificity=48.5%, AUC=0.743), VHI-10(cutoff=13.0, sensitivity=77.9%, specificity=89.7%, AUC=0.893), Grade(cutoff=2.0, sensitivity=70.6%, specificity=69.1%, AUC=0.728)와 OS(cutoff=51.0, sensitivity=86.8%, specificity=66.2%, AUC=0.833)로 분석되었다.

고 찰

본 연구에서는 후두미세수술을 실시한 성대용종 환자들의 술 전, 술 후 음성을 AVQI와 GRBAS, CAPE-V를 통해서 비교하였다. 음향학적 변수들과 청지각적 변수들은 유의한 긍정적 변화가 나타났으며 음성의 회복을 예측할 수 있었다. 또한, 음향학적 변수들은 청지각적 평가들과는 비교적 높은 상관관계를 보인 반면 환자 스스로 평가하는 설문지와는 낮은 상관관계를 보였다. 이는 음성 회복의 정도와 환자 스스로 느끼는 회복의 정도가 다를 것을 의미한다. 음성 설문지의 경우 환자들이 주의 깊게 작성하지 않은 것에도 연관이 있는 것으로 사료된다. 임상현장에서 음성분석을 진행할 경우 매우 나쁜 음성의 MDVP 분석이 정확하게 이루어지지 않거나, 모음연장발성과 문장 읽기(발화주파수) 등의 음성이 상이하게 분석되는 경우가 많다. 이러한 경우 숙련된 임상가의 경우

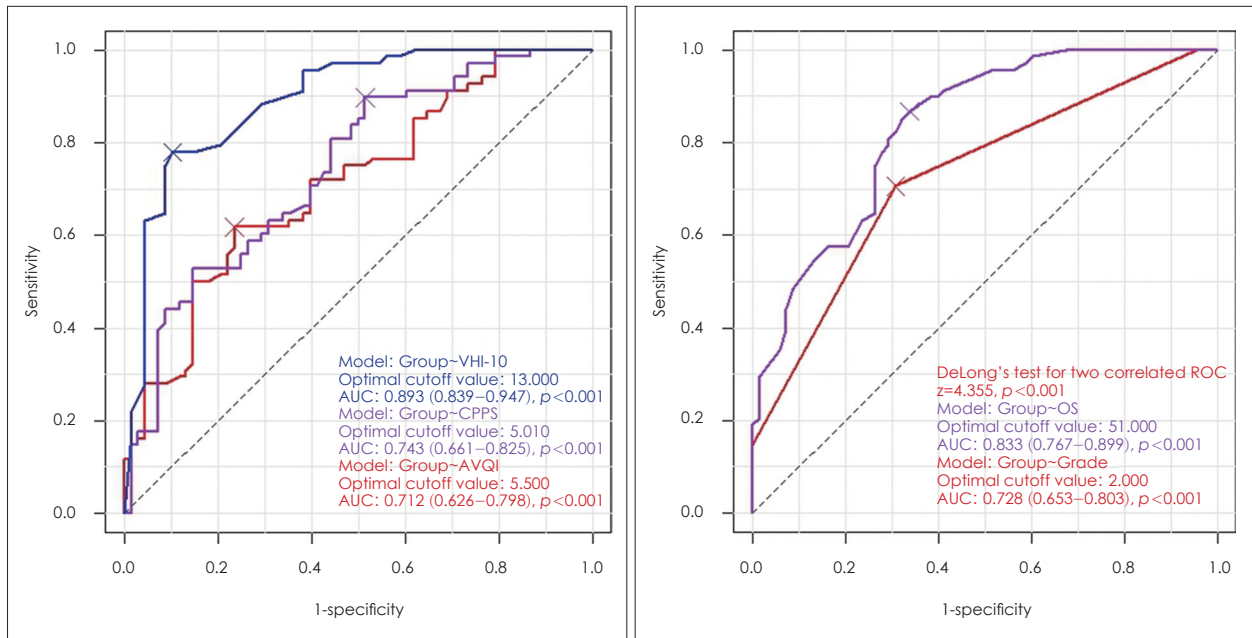


Fig. 5. ROC curve analysis for discrimination of voice recovery between pre/post operative evaluation. AUC: area under the curve, AVQI: acoustic voice quality index, OS: overall severity, ROC: receiver operating characteristic, VHI-10: voice handicap index-10.

재검사를 통해서 음성평가의 신뢰도를 높이지만 그렇지 못한 경우에는 다른 분석 결과가 도출되는 경우도 있다. AVQI는 음성평가의 과업에 대한 음성을 하나의 음성파일로 만들어 무성구간 제거 후 자동 분석된다. 이러한 AVQI는 청지각적 평가와도 높은 상관관계를 나타내었고 후두미세수술 후 음성 회복에 대해서도 확인이 가능한 요인이었다.

후두미세수술 전, 후의 음성분석에서 통계적으로 유의한 변수는 AVQI, CPPS, G, OS, VHI-10였다. 5개의 유의미한 평가 변수 중 음향학적 측정 관련 변수인 AVQI, CPPS는 후두미세수술 후 음성이 개선되는 수치로 변화했다. 정상적인 성대 진동 및 성문폐쇄를 방해하는 성대용종이 제거되어 음향학적 측정의 결과가 호전된 것으로 나타났다. 청지각적 평가 관련 변수인 G, OS와 환자 스스로 본인의 음성을 확인하는 음성 관련 설문지(VHI-10)에서도 후두미세수술 후 평가 점수가 감소하여 평가자 및 환자 본인도 음성이 호전되었다고 판단한 것으로 생각한다.

성대용종은 Titze⁴⁾가 제시한 음성신호의 분류법에 따라 Type 1, 2, 3으로 나뉘질 수 있다. Type 1은 물리적 변동이 거의 없는 주기적인 신호, Type 2는 변조된 주파수 혹은 저조파(sub-harmonic)를 포함한 신호, Type 3은 주기성을 확인할 수 없거나 혹은 혼돈적 신호(chaos signal)라고 하였다. Type 2, 3에 해당하는 음성들은 섭동 분석 방식으로는 한계가 있다고 하였다.^{17,18)} 본 연구에서는 AVQI를 통해서 모든 유형의 음성신호를 분석하여 정량화 할 수 있으며 모음연장발성과 문장 읽기 과업을 한 번에 분석하였다. 개별 과업을 분석한 결

과보다 실제 환자의 음성 특성을 더욱 반영할 수 있는 결과라고 생각된다.

AVQI는 CPPS, G, OS와 상관관계가 있는 것으로 확인된 반면 VHI-10과는 상관관계가 없는 것으로 확인되었다. Shin 등¹⁹⁾과 Uloza 등²⁰⁾의 연구에서 후두미세수술 후 GBBAS, jitter, shimmer, noise to harmony ratio 등이 유의하게 개선되었다고 보고하였다. 또한, 객관적 측정(MDVP, jitter, shimmer)이 주관적 평가(GRBAS)와 높은 상관관계가 있고 jitter, shimmer를 통해서 치료의 호전을 확인할 수 있었다. AVQI와 청지각적 평가와의 상관에 대한 연구는 지속적으로 보고되었다. 연구 초기에는 GRBAS에 대한 보고가 많았으며 이후 CAPE-V도 함께 보고되었다. AVQI와 GRBAS의 상관관계는 영어권 0.794~0.868, 독일어권 0.780~0.858, 프랑스어 0.781, 일본어 0.828, 한국어 0.911, 이탈리아에서는 G, B와 0.684로 보고되었다. AVQI와 OS의 상관관계는 독일어권 0.852~0.876, 한국어 0.924로 보고되었다. 5~11점의 동등간격척도(Equal-Appearing Interval Scale)와 다른 일반적인 측정 변수(pitch, intensity, F0, jitter, shimmer, HNR, spectrum, subglottic pressure, maximum phonation time, soft phonation index etc.) 간의 상관관계²¹⁻²⁶⁾는 0.560~0.880, 분류 정확성(classification accuracy)^{25,27-29)}은 49.9~86.0%로 보고되었다. 본 연구에서도 선행연구들의 결과와 같이 음향학적 측정 변수들과 청지각적 평가 사이에 높은 상관관계가 확인되었다. 반면에, 다른 변수들과 VHI-10과의 상관관계는 상대적으로 낮았으며, 술 전에는 CPPS($r = -0.294$), G($r = 0.347$), OS($r = 0.345$)의

상관관계, 술 후에는 CPPS($r=-0.242$), OS($r=0.305$)의 상관관계가 있었다. 대부분 낮은 상관관계 혹은 상관관계가 없는 것으로 나타났으며 이는 VHI와 음질 측정 변수들 간에는 상관관계가 없다고 보고한 선행연구³⁰⁾의 결과와 일치한다. VHI 설문지는 주관적이며 음성문제를 가진 환자의 진단에 있어서 그 자체는 중요하지 않다는 것을 시사하였다. 설문지 자체는 음성장애 환자들의 삶의 질을 확인하기 위한 절차로 사용될 수 있다고 하였다.

유경형 성대용종의 크기에 따른 술 전, 술 후의 음성은 모든 크기에서 호전된 것으로 나타났고, 통계적으로 유의한 차이는 small, medium의 크기에서만 확인되었다. Large의 크기에서는 유의한 차이가 나타나지 않은 것은 대상자의 수가 적은 것(4명)으로 사료된다. 범발성 성대용종은 넓은 범위에 발생하여 성대용종 크기에 따른 분류를 하지 않았다. 선행연구에서 성대용종의 색깔, 유형, 크기는 음성치료에 대한 반응율과 유의미한 상관이 없는 것으로 보고하였다.³¹⁾ Small 집단에서는 66명 중 37명(56.1%), medium 집단에서는 14명 중 2명(14.3%), large 집단에서는 12명 중 1명(8.3%)이 음성치료에 반응을 보였다. 효과적인 음성치료를 위한 지표는 여성과 small 크기였고, 유경형보다는 범발성 성대용종에서 좋은 효과가 보고되었다. Jeong 등¹⁴⁾은 여성, small 크기, 증상의 지속기간이 짧은 환자들이 보존치료(음성치료) 등을 통해서 개선될 가능성이 더 높다고 보고하였다. 이러한 결과들을 바탕으로 후두미세수술, 음성치료를 고려한다면 환자의 만족도는 더욱 높아질 것으로 사료된다.

후두미세수술 전후 음성에 대한 최적의 절단값 및 곡선 아래 면적을 구하여 음성의 회복 여부에 대한 참고값을 확인하였다. 수술 전후 AVQI 경계의 기준값은 5.5였으며 선행연구들에서 나타난 정상과 병리적 음성 간 경계의 기준값은 2.97~3.66 사이에 분포하였다.⁷⁻¹¹⁾ 본 연구에서는 선행연구의 정상 기준값보다 높은 값이 측정되었고 후두미세수술 후의 음성이라도 하더라도 완벽하게 정상음성으로 산출되지는 않았음을 의미한다. 술 후에도 음성위생, 음성관리를 잘 관리하여 정상음성으로 회복할 수 있도록 해야 한다. 본 연구의 AVQI에 대한 AUC(0.712)는 선행연구들에 비해서 낮게 측정되었지만 Swets³²⁾에 의하면 중증도의 정확한($0.7 < \text{AUC} < 0.9$) 검사로 분류될 수 있다. 또한 VHI-10의 AUC가 가장 크게 나타났으며, 이것은 환자의 음성 회복에 대한 판단은 환자 스스로 체감하는 호전의 정도도 참고되어야 함을 의미한다. 평가자가 판단한 음성장애의 정도와 환자 스스로 생각하는 본인의 음성장애 정도에는 차이가 있지만, 환자의 불만과 불편감을 알게 된다면 적절한 치료의 방향을 수립할 수 있어 추가 치료의 효과를 증대시킬 수 있다.³³⁾ 또한 음성문제를 자각

한다는 것은 건강한 삶의 질을 위해서도 중요하며, 개선된 목소리에 대한 만족도를 통해 치료 효과를 알 수 있다.³⁴⁾

임상현장에서는 음성 회복의 정도를 확인하기 위해서 다양한 평가 방법들이 사용되고 있다. 후두내시경, 후두스트로보스코피, 전기성문파형검사(electroglottography) 등과 같은 다른 검사들로도 확인할 수 있지만 비침습적이고 환자에게 주는 불편감이 없어 음향학적 음성평가가 주로 사용되고 있다. 후두의 병변, 음성장애의 원인에 따라서 모음연장발성과 문장 읽기의 결과가 상이하게 나타나는 경우도 있다. 이런 경우 임상가의 주관, 경험에 의해서 판단되며 평가자에 따라서 다르게 판단될 수 있다. 섭동 분석은 매우 비주거적인 음성 신호를 가진 경우에는 정확한 분석이 되지 않으며, 이 부분을 보완하기 위해 캡스트럼 분석을 실시하지만 쉼, 무성음 구간에 따라 CPPS의 값이 달라지는 제한점이 있다. CPPS는 이 변수들에 대한 영향을 많이 받아서 호흡지지가 부족한 대상자들은 호흡지지가 충분한 대상자들에 비해 낮은 CPPS가 측정된다.

AVQI는 동일한 모음연장 길이 및 문장 읽기, 여기에 추가적으로 문장 읽기에서 쉼, 무성음 구간을 Praat 스크립트로 제거하여 유성음 구간만을 분석한다. 모음연장의 길이를 연구자의 의도대로 설정할 수 있으며, 보완할 부분이 있다고 판단되면 새로운 회귀식에 대한 검증도 진행할 수 있다. 이전 버전의 AVQI 회귀식은 SpeechTool을 통해서 CPPS를 측정하고 나머지 변수들은 Praat을 통해서 측정하였다. 본 연구에서 사용된 AVQI 회귀식은 Praat에서 CPPS를 포함한 모든 변수를 측정하도록 설계되어 있다. 최대한 동일한 조건의 과업에서 얻어진 음성샘플을 분석하여 음성장애의 정량화를 할 수 있으며, 동일한 모음연장 구간과 문장 읽기를 통일한다면 동일 언어권 내에서는 음성장애의 정도를 비교하는 것이 가능할 것이다. 음향학적, 청지각적 평가, 음성 설문지를 통해서 후두미세수술 전후의 음성을 분류할 확률이 71.2~89.3%이다. 음향학적 평가 단독보다는 다양한 평가 변수들을 함께 분석하여 판단하면 음성 회복에 대한 예측을 높일 수 있으며, 본 연구를 통해서 새롭게 제시된 AVQI를 포함한 음향학적, 청지각적 평가들이 음성장애 선별 도구로 충분히 활용될 수 있는 가능성을 확인하였다고 생각한다.

본 연구의 제한점은 AVQI 측정을 위한 회귀식은 지속적으로 개발되고 업데이트되고 있어 다양하다는 것이다. 본 연구에서 적용한 AVQI의 버전은 2.02지만 현재 3.01도 소개되어 있다.^{35,36)} 병리적 음성에 대한 정량화 및 감별을 위해서는 새롭게 제시된 회귀식도 지속적으로 적용하여 정확도를 높여야 할 것이다. 최근 발표된 연구들은 대부분 2.02 버전을 적용하였고 3.01에 대한 검증도 진행하고 있다. 선행연구들은

AVQI에 대한 타당도를 검증과 AVQI 회귀식의 버전 업데이트하는 보고가 많았으며, 음성장애 환자에 대한 치료 전후의 음성 회복 정도를 확인하기 위한 연구는 보고되지 않았다. 추후 연구에는 동일한 음성에 대해서 회귀식에 따른 값들을 비교하여야 하며, 성대용종에 대한 사례를 모아서 유형에 따른 비교, 성대용종 크기에 따른 비교가 이루어져야 할 것이다. 후두미세수술 뿐만 아니라 음성치료, 성대 내 주입술 등 다양한 치료에 대한 효과를 검증하기 위해서도 AVQI 측정이 이루어져야 할 것이다.

REFERENCES

- 1) Johns MM, Garrett CG, Hwang J, Ossoff RH, Courey MS. Quality-of-life outcomes following laryngeal endoscopic surgery for nonneoplastic vocal fold lesions. *Ann Otol Rhinol Laryngol* 2004; 113(8):597-601.
- 2) Hochman II, Zeitels SM. Phonomicrosurgical management of vocal fold polyps: the subepithelial microflap resection technique. *J Voice* 2000;14(1):112-8.
- 3) Makeieff M, Barbotte E, Giovanni A, Guerrier B. Acoustic and aerodynamic measurement of speech production after supracricoid partial laryngectomy. *Laryngoscope* 2005;115(3):546-51.
- 4) Titze IR. Workshop on Acoustic Voice Analysis: Summary statement. Denver(CO): National Center for Voice and Speech;1995. p.1-36.
- 5) Maryn Y, De Bodt M, Roy N. The acoustic voice quality index: toward improved treatment outcomes assessment in voice disorders. *J Commun Disord* 2010;43(3):161-74.
- 6) Maryn Y, Corthals P, Van Cauwenberge P, Roy N, De Bodt M. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice* 2010;24(5):540-55.
- 7) Barsties B, Maryn Y. [The acoustic voice quality index. toward expanded measurement of dysphonia severity in German subjects]. *HNO* 2012;60(8):715-20.
- 8) Maryn Y, De Bodt M, Barsties B, Roy N. The value of the acoustic voice quality index as a measure of dysphonia severity in subjects speaking different languages. *Eur Arch Otorhinolaryngol* 2014; 271(6):1609-19.
- 9) Maryn Y, Kim HT, Kim J. Auditory-perceptual and acoustic methods in measuring dysphonia severity of Korean speech. *J Voice* 2016; 30(5):587-94.
- 10) Hosokawa K, Barsties B, Iwahashi T, Iwahashi M, Kato C, Iwaki S, et al. Validation of the acoustic voice quality index in the Japanese language. *J Voice* 2017;31(2):260.e1-9.
- 11) Uloza V, Petrauskas T, Padervinskis E, Ulozaitė N, Barsties B, Maryn Y. Validation of the acoustic voice quality index in the lithuanian language. *J Voice* 2017;31(2):257.e1-11.
- 12) Ingrisano DR, Perry CK, Jepson KR. Environmental noise: a threat to automatic voice analysis. *Am J Speech Lang Pathol* 1998;7(1): 91-6.
- 13) Deliyski DD, Shaw HS, Evans MK. Adverse effects of environmental noise on acoustic voice quality measurements. *J Voice* 2005;19(1): 15-28.
- 14) Jeong WJ, Lee SJ, Lee WY, Chang H, Ahn SH. Conservative management for vocal fold polyps. *JAMA Otolaryngol Head Neck Surg* 2014;140(5):448-52.
- 15) Kleinsasser O. Pathogenesis of vocal cord polyps. *Ann Otol Rhinol Laryngol* 1982;91(4 Pt 1):378-81.
- 16) Núñez-Batalla F, Díaz-Fresno E, Álvarez-Fernández A, Muñoz Cordero G, Llorente Pendás JL. [Application of the acoustic voice quality index for objective measurement of dysphonia severity]. *Acta Otorrinolaringol Esp* 2017;68(4):204-11.
- 17) Carding PN, Steen IN, Webb A, MacKenzie K, Deary IJ, Wilson JA. The reliability and sensitivity to change of acoustic measures of voice quality. *Clin Otolaryngol Allied Sci* 2004;29(5):538-44.
- 18) Choi SH, Choi CH. [The utility of perturbation, non-linear dynamic, and cepstrum measures of dysphonia according to signal typing]. *Phonetics and Speech Sciences* 2014;6(3):63-72.
- 19) Shin YS, Chang JW, Yang SM, Wu HW, Cho MH, Kim CH. Persistent dysphonia after laryngomicrosurgery for benign vocal fold disease. *Clin Exp Otorhinolaryngol* 2013;6(3):166-70.
- 20) Uloza V, Saferis V, Uloziene I. Perceptual and acoustic assessment of voice pathology and the efficacy of endolaryngeal phonomicrosurgery. *J Voice* 2005;19(1):138-45.
- 21) Eskenazi L, Childers DG, Hicks DM. Acoustic correlates of vocal quality. *J Speech Hear Res* 1990;33(2):298-306.
- 22) Wolfe V, Fitch J, Cornell R. Acoustic prediction of severity in commonly occurring voice problems. *J Speech Hear Res* 1995;38(2):273-9.
- 23) Wolfe V, Fitch J, Martin D. Acoustic measures of dysphonic severity across and within voice types. *Folia Phoniatr Logop* 1997;49(6): 292-9.
- 24) Piccirillo JF, Painter C, Fuller D, Haiduk A, Fredrickson JM. Assessment of two objective voice function indices. *Ann Otol Rhinol Laryngol* 1998;107(5 Pt 1):396-400.
- 25) Yu P, Ouaknine M, Revis J, Giovanni A. Objective voice analysis for dysphonic patients: a multiparametric protocol including acoustic and aerodynamic measurements. *J Voice* 2001;15(4):529-42.
- 26) Bhuta T, Patrick L, Garnett JD. Perceptual evaluation of voice quality and its correlation with acoustic measurements. *J Voice* 2004;18(3): 299-304.
- 27) Giovanni A, Robert D, Estublier N, Teston B, Zanaret M, Cannoni M. Objective evaluation of dysphonia: preliminary results of a device allowing simultaneous acoustic and aerodynamic measurements. *Folia Phoniatr Logop* 1996;48(4):175-85.
- 28) Wuys FL, De Bodt MS, Molenberghs G, Remacle M, Heylen L, Millet B, et al. The dysphonia severity index: an objective measure of vocal quality based on a multiparameter approach. *J Speech Lang Hear Res* 2000;43(3):796-809.
- 29) Ma EP, Yiu EM. Multiparametric evaluation of dysphonic severity. *J Voice* 2006;20(3):380-90.
- 30) Pinarbasli MO, Kaya E, Ozudogru E, Gurbuz MK, Colak E, Aksoy MA, et al. Acoustic analysis of soccer fans in acute phonotrauma after the match. *J Voice* 2017 Nov 13 [Epub]. available from: <https://doi.org/10.1016/j.jvoice.2017.10.004>.
- 31) Lee YS, Lee DH, Jeong GE, Kim JW, Roh JL, Choi SH, et al. Treatment efficacy of voice therapy for vocal fold polyps and factors predictive of its efficacy. *J Voice* 2017;31(1):120.e9-13.
- 32) Swets JA. Measuring the accuracy of diagnostic systems. *Science* 1988;240(4857):1285-93.
- 33) Kim JO, Choi HS. [Comparison of clinicians' perceptual evaluations and patients' subjective evaluations of voice disorders]. *Commun Sci Disord* 2009;14(2):223-35.
- 34) Lee OB, Kwon SB. [Development of a Korean version of the self awareness index for voice problems: preliminary study]. *Journal of Speech and Hearing Disorders* 2011;20(2):107-23.
- 35) Hosokawa K, v Latoszek BB, Iwahashi T, Iwahashi M, Iwaki S, Kato C, et al. The acoustic voice quality index version 03.01 for the Japanese-speaking population. *J Voice* 2017 Nov 15 [Epub]. available from: <https://doi.org/10.1016/j.jvoice.2017.10.003>.
- 36) Barsties B, Maryn Y. External validation of the acoustic voice quality index version 03.01 with extended representativity. *Ann Otol Rhinol Laryngol* 2016;125(7):571-83.