Analysis of Voice Changes in Early-Stage Parkinson's Disease with AVQI and ABI: A Follow-up Study

Rachel B. Convey, Anne-Maria Laukkanen, Sari Ylinen, and Nelly Penttilä, *Tampere, Finland*

Summary: Objectives. The purpose of this pilot study was to examine voice quality changes in individuals with early-stage Parkinson's disease (PD) utilizing the Acoustic Voice Quality Index (AVQI) and Acoustic Breathiness Index (ABI) over approximately a 1-year period.

Study design. Follow-up study.

Methods. Baseline and follow-up data were gathered from the PDSTUlong speech corpus. The data for both time points included: speaker background information, sustained vowels, reading samples, and measures of PD severity (Hoehn and Yahr scores and Unified Parkinson's Disease Rating Scale III scores [UPDRS-III]). All speakers (N = 12) were native Finnish speakers. AVQIv03.01 and ABI analysis were completed in VOXplot v2. 0.1. Changes in AVQI and ABI scores between baseline and follow-up were examined via causal analysis. Further, AVQI and ABI were analyzed in relation to measures of PD severity.

Results. Baseline mean AVQI score was 1.79 (range 0.14–4.83, SD = 1.60), whereas follow-up mean AVQI score was 2.25 (range 0.55–4.53, SD = 1.36). Baseline mean ABI score, in turn, was 2.92 (range 1–27 – 5.31, SD = 1.57), whereas follow-up mean ABI score was 3.42 (range 1.40–5.40, SD = 1.38). A significant difference was found between baseline and follow-up measures for both AVQI ($Z = -2.002$, $P = 0.045$) and ABI ($Z = -2.197$, $P = 0.028$). A significant difference in smoothed cepstral peak prominence ($Z = -2.118$, $P = 0.034$) and harmonics-to-noise ratio ($Z = -1.961$, $P = 0.050$) was also found between the two measurement periods. Change in AVQI and ABI were not correlated with the change in measures of PD severity.

Conclusion. Over approximately 1-year, a statistical change was observed in AVQI and ABI scores, even in such a small dataset. The specific qualities of breathiness and hoarseness showed the most significant progression. Changes in voice quality were more prominent in ABI analysis.

Key Words: Acoustic analysis–Early-stage Parkinson's disease–AVQI–ABI.

INTRODUCTION

Individuals with Parkinson's disease (PD) experience changes in voice, specifically in fundamental frequency, approximately 5 years prior to receiving a medical diagnosis.[1] Though early voice changes have perceptual correlates, clinical ears may not be sensitive enough to detect these changes, making acoustic measures the best means for objective analysis.[2] Knowledge of how these changes are reflected in acoustic measures would enable us to identify the voice biomarkers indicative of PD more precisely, aiding in early disease detection. The aim of this pilot study is to analyze how early voice changes in PD are reflected in the acoustic voice quality index (AVQI) and the acoustic breathiness index (ABI) over the period of approximately 1-year.

As PD progresses, around 90% of individuals will be diagnosed with hypokinetic dysarthria.[3,4] Breathy, rough and asthenic (weak) vocal quality, increased fundamental frequency and monoloudness characterize this motor speech disorder.[5,6] These changes in voice quality result from alterations in the vocal mechanism, including incomplete glottal closure, phase asymmetry, vocal fold bowing, and increased glottal open time during speech.[2] It is suspected that these voice changes arise secondary to the decreased level of dopamine associated with PD.[2]

In the clinical setting, Speech-Language Pathologists (SLP) diagnose hypokinetic dysarthria primarily based on perceptual observations. Rating scales, like the Consensus Auditory-Perceptual Evaluation of Voice (CAPE-V) or GRBAS scale (parameters rated: grade, roughness, breathiness, asthenia, strain), are employed in analysis of voice quality. Measures of voice quality may be a facet employed in the tracking of communication changes in PD.[7]

Acoustic measures offer an objective means of analysis, not influenced by clinical experience or other human factors. Voice changes in PD that can be measured acoustically include increased shimmer, increased jitter, reduced harmonics-to-noise ratio (HNR) and changes in fundamental frequency.[2,8] Generally, changes in voice are more common in early PD compared to changes in articulation precision and fluency.[9] Acoustic measurement of voice could support earlier diagnostics.[10] Given the greater ability of acoustic measures to detect the early voice changes in individuals with PD, this study will focus solely on acoustic measures, though we acknowledge that

Accepted for publication May 18, 2024.

RBC was supported to conduct this research by a grant from Finnish National Agency for Education, Opetushallitus Utbildningsstyrelsen, grant number: OPH-3748-2022 and by a grant from the Päivikki and Sakari Sohlberg Foundation.

From the Faculty of Social Sciences, Tampere University, Tampere, Finland.

Address correspondence and reprint requests to Rachel B. Convey, Faculty of Social Sciences, Tampere University, Kalevantie 5, 33100 Tampere, Finland. E-mail: rachel.convey@tuni.fi

Journal of Voice, Vol xx, No xx, pp. xxx–xxx

0892-1997

https://doi.org/10.1016/j.jvoice.2024.05.009

perceptual analysis of voice is often completed in conjunction with acoustic analysis in the clinical setting.

Rusz and colleagues (2011) attempted to identify acoustic measures that could distinguish early untreated PD speakers from healthy speakers. Of the early untreated PD speakers in the study, 78% showed some type of voice impairment. Broadly, most of the untreated PD speakers presented with disordered prosody in comparison to disordered phonation or articulation.[11] In differentiating early untreated PD speakers from healthy controls, the measure of fundamental frequency in monologue and emotional sentence reading was found the be the most sensitive.[11] However, measures of jitter, shimmer (local), HNR and noise-to-harmonics ratio, all measures of perturbation, were also sufficient in distinguishing early untreated PD from healthy control speakers.

Voice features are capable of distinguishing patients with PD from healthy speakers,[2] and differences have also been observed among different disease subtypes.[12] However, from a clinical perspective, it is important to examine the progression of the disease and the relationship between voice symptoms and disease severity. While there are known voice changes that occur in PD, further differentiation has been found between the early stages of early-onset PD (EOPD) and late-onset PD (LOPD). Decreased voice quality was found to classify speakers with LOPD, while weaker inspirations classified speakers with EOPD.[12] Both groups exhibited imprecise articulation, monopitch (SD F0), and monoloudness (SD dB).[12]

Multiple studies have illustrated the complex relationship between changes in voice and disease progression in PD. A recent study analyzed the ability of the voice signal to predict PD severity as measured by the Unified Parkinson's Disease Rating Scale III (UPDRS-III).[13] Under the study conditions, results showed that analysis of the vowel /a/ was the best predictor of UPDRS-III score 180 minutes following levodopa intake, advancing the possibility of accurately accessing disease stage outside of the clinical setting.[13] Past studies have found that voice features associated with hypokinetic dysarthria, including harsh vocal quality, decreased intensity, and increased jitter, do not progress along with disease severity.[14] Conducted over the course of 1-year, a follow-up study by Skodda and colleagues (2013) shows similar findings. Changes in shimmer, jitter, noise-to-harmonics ratio, decreased speech rate, and deterioration of articulation were observed, while the measure of PD severity (UPDRS) remained relatively unchanged across the cohort of speakers.[7] However, other symptoms including breathy vocal quality and reduced loudness became more severe as PD progresses into the later stages.[14] As illustrated through the mixed results presented above, the relationship between changes in voice and PD stage remains complex and not clearly defined.

Perceptual measures are applied in the clinical work of SLPs for analysis of voice quality. These measures are subjective in nature and for that, we need research on the functionality of various acoustic measures for application in the clinical setting. This study focuses specifically on two acoustic measures: the Acoustic Voice Quality Index (AVQI) and the Acoustic Breathiness Index (ABI). These measures are relatively new and have been recently validate for use with speakers of the Finnish language.[15,16] AVQI was first introduced in 2010.[17,18] This multiparametric acoustic measure can distinguish a healthy voice from a dysphonic voice and is unique in its consideration of both a sustained vowel and continuous speech sample.[17] Each language has a unique AVQI cutoff score. In validating AVQI version 03.01 for Finnish, the cutoff score is 1.83 for AVQI version 03.01.[16] This cutoff score separates dysphonic from nondysphonic voices, and thus the voice of any speaker whose AVQI score is greater than 1.83 is considered dysphonic.

AVQI is composed of the following parameters: HNR, shimmer (dB), shimmer (%), smoothed cepstral peak prominence (CPPs), long-term average spectrum (LTAS) tilt, and slope.[17] AVQI is constructed from the subparameters by the following equation[17]:

$$\begin{aligned} AVQI = (3.295 &- (0.111 \times CPPs) - (0.073 \times HNR) \\ &- (0.213 \times shimmer\ local) \\ &+ (2.789 \times shimmer\ local\ dB) - (0.032 \times slope) \\ &+ (0.077 \times tilt) \times 2.571 \end{aligned}$$

AVQI analysis can be completed in either Praat or VOXplot free software.[19,20] Analysis in Praat requires the use of the AVQIv03.01 script. For analysis with either software, the cut 3-second long sustained /a/ speech sample and continuous speech sample (first 33 syllables of the passage "Pohjantuuli ja Aurinko" (The Northwind and the Sun) for Finnish language analysis) are inputted. Following this step, one is able to run the analysis and receive the AVQI output composed of the AVQI score and subparameter values.[16]

AVQI has been applied in analyzing voice among speakers of varying populations. In speakers with PD, specifically, several studies have used this measure to analyze voice quality. Studies conducted by Boutsen et al[21] and Moya-Galé et al[22] utilized AVQI to analyze voice changes in speakers with PD secondary to SpeakOUT!® and Lee Silverman Voice Treatment (LSVT®LOUD) intervention, respectively. Results from both studies illustrated improved voice quality in speakers, as measured by AVQI, following therapeutic intervention. Further, Convey et al[23] examined the relationship between AVQI analysis and perceptual analysis of voice quality completed with the GRBASI scale in speakers with PD. In this study, AVQI was found to be significantly correlated with the perceptual analysis but not with the disease stage as measured by the Hoehn and Yahr scale. Further research on the relationship between PD stage and changes in AVQI has not yet been conducted. Additionally, AVQI has not yet been used to analyze early voice changes in PD.

ABI was developed in 2017 and validated for use with Finnish language speakers in 2023.[15,24] Similar to AVQI, ABI cutoff score for each language is unique. For the Finnish language, the cutoff score is 2.68.[15] This cutoff score separates breathy from nonbreathy voices, the voice of any speaker whose ABI score is greater than 2.68 is considered breathy. ABI analyzes breathy voice quality, a perceptual phenomenon in part resulting from incomplete glottal closure and decreased subglottal pressure.[2,8]

This multiparametric acoustic measure is composed of the following parameters: CPPs, period standard deviation (PSD), jitter local (Jit), glottal-to-noise excitation ratio (GNE), high-frequency noise of 6000 Hz (Hfno), amplitude difference between the first two harmonics (H1-H2), harmonics-to-noise ratio of Dejonckere (HNR-D), and shimmer (dB and %).[24] ABI is represented by the following equation[24]:

$$\begin{aligned} ABI = (&5.0447730915 - [0.172 \times CPPs] - [0.193 \times Jit] \\ &- [1.283 \times GNE\,\max - 4500Hz] \\ &- [0.396 \times Hfno - 6000Hz] + [0.01 \times HNR - D] \\ &+ [0.017 \times H1 - H2] + [1.473 \times Shim\,dB] \\ &- [0.088 \times Shim] - [68.295 \times PSD]) \\ &\times 2.9257400394 \end{aligned}$$

ABI analysis can be completed with VOXplot free software and is reliant on both a sustained vowel (3-second long sustained /a/) and continuous speech sample (first 33 syllables of the passage "Pohjantuuli ja Aurinko" (The Northwind and the Sun) for Finnish language analysis).[15,20] ABI has not yet been applied to analysis of PD voice.

The purpose of this pilot study is to analyze voice changes in speakers with early-stage PD utilizing two acoustic measures, AVQI and ABI, over the period of approximately 1-year. For this study, early-stage PD is defined as PD that has been diagnosed within 2 years. Through this study, we aim to expand our knowledge of how early voice changes in PD are reflected in objective, multiparametric acoustic measures. We address the following research questions: (1) Between the two measurement points (baseline and the approximate 1-year follow-up), does voice quality deteriorate as measured by AVQI? (2) Between the two measurement points (baseline and the approximate 1-year follow-up), does voice quality deteriorate as measured by ABI? (3) Are there correlations between the change in AVQI and ABI scores and the change in measures of PD stage (Hoehn and Yahr and UPDRS-III) over the measurement period? We hypothesize that (1) poorer voice quality, illustrated by increased AVQI scores, will be observed in the follow-up period compared to the baseline, (2) increased breathiness, as indicated by increased ABI scores, will be observed in the follow-up period compared to the baseline and (3) no correlation will be found between the change in AVQI and ABI scores and the change in ratings of PD stage.

METHODS

Ethical permission

The speech corpus used in this study belongs to the project titled "*Akustisen ja audiovisuaalisen signaalin käyttö ihmisen terveydentilan analyysissa ja automaattisessa luokittamisessa*" *(The use of acoustic and audiovisual signals in the analysis and automatic classification of human health status),* referred to as AAVISTUS. This project is led by docent Nelly Penttilä and run in collaboration with the Wellbeing services of the county of Pirkanmaa, Finland. Ethical permission for AAVISTUS was assessed by the Regional Ethics Committee of the Expert Responsibility area of Tampere University Hospital in November 2021 (ETL code R21122). Research permits for the project were granted to Tampere University Hospital (Tays) and the city of Tampere in December 2021. The research is conducted at the University of Tampere.

Speech corpus

PD speakers are recruited for AAVISTUS by participating neurologists. When recruited, speakers may be suspected of having PD or have been diagnosed with PD for less than 2 years. However, all speakers included in the Parkinson's Disease Speech corpus of Tampere University long (PDSTUlong) for which follow-up data has been collected have a confirmed diagnosis of PD. Those speakers who were suspected of having PD, but who were not diagnosed with the disease were excluded from the speech corpus.[25] Speakers in PDSTUlong met the following inclusion criteria: suspected PD by neurologist or 0–2 years since Parkinson's diagnosis, diagnosis of PD during the follow-up period, Finnish as the native language, no history of head or neck cancers, no prior neurological cerebrovascular events or diagnosed diseases, no previously diagnosed voice disorders, or speech, language developmental disorders. They are notified of the AAVISTUS study details, purpose, and their right to exit the study at any point at the beginning of their participation.

PDSTUlong contains acoustic, video, and speaker background data for two measurement points: baseline and follow-up. Data collection is ongoing and includes audio recordings from eight different speech tasks, audiovisual recordings of facial expressions during data collection tasks, self-assessments of experienced symptoms related to voice (Voice Handicap Index 9), swallowing (eat10), depression (pdq9), and anxiety (gad7), health, treatment, and imaging information regarding the disease (patient information system) like H&Y rating and UPDRS-score.

Recordings were made in a quiet setting. A 4066 omnidirectional headset microphone (DPA microphones, Denmark) was utilized. The headset microphone was placed at a 45-degree angle from the corner of the mouth at a distance of 4 cm. Recordings were made in WAV (Waveform Audio File) format with PRAAT software[26] and Focusrite audio interface.

TABLE 1.
Speaker Background Information

Speaker ID	Sex (M/F)	Age at PD diagnosis (years)	Age at baseline (years)	Time between baseline and follow-up (weeks)	Speaker PD medication status (Yes/No)		Speaker vocation
					Baseline	Follow-up	
1	F	71.67	72.75	53.29	Yes	Yes	Healthcare
2	F	66.75	67.17	54.57	Yes	Yes	Finance
3	M	64.75	64.67	54.57	Yes	Yes	Industry
4	F	79.83	79.83	59.57	Yes	Yes	Healthcare
5	F	69.75	69.67	58.57	Yes	Yes	Finance
6	F	49.58	49.75	79.29	Yes	Yes	Education
7	M	66.75	67.08	70.29	Yes	Yes	No reported profession
8	M	62.42	62.83	69.71	Yes	Yes	Education
9	F	84.33	85.00	58.29	Yes	Yes	Healthcare
10	F	69.17	69.33	55.43	Yes	No	Service industry
11	F	77.50	77.75	49.71	Yes	Yes	Healthcare, Education
12	F	73.58	74.67	53	Yes	Data not available	Healthcare
	Average	69.67	70.04	59.69			
	Minimum	49.58	49.75	49.71			
	Maximum	84.33	85.00	79.29			
	SD	9.03	9.14	8.82			

Speaker selection and background data

In total, there are 24 speakers in PDSTUlong with PD. For this study, 12 were selected based on the following criteria: confirmed diagnosis of PD, measurement points obtained for baseline and follow-up. The average duration between these measurement points was 59.69 weeks. The medical background data, gathered from the speaker's medical record, includes speaker sex, birthdate, date of PD diagnosis, Hoehn and Yahr scores, UPDRS-III score and prescribed medications at the time of data collection and vocational history. The speech samples collected for this study include a prolongation of the vowel /a/ and a normal reading of the passage "Pohjantuuli ja Aurinko" (The Northwind and the Sun).

Speaker background data are presented in Table 1. Three speakers are male, nine are female. Average speaker age at time of PD diagnosis was 69.67 years and ranged from 49.58–84.33 years (SD = 9.03). The average time between the two measurement points (baseline to follow-up) was 59.69 weeks (range = 49.71–79.29 weeks, SD = 8.82).

AVQI and ABI analysis

All ABI and AVQI analyses completed using VOXplot v2.0.1.[20] VOXplot is a free software designed specifically for voice quality analysis. When using VOXplot, one is given the option to specify the language of analysis. For the purposes of this study, the Finnish language was specified. The same sustained vowel and continuous speech samples are required for AVQI and ABI analysis. The sustained vowel sample consists of the center 3 seconds of the sustained vowel /a/. The continuous speech sample consists of the first 33 syllables of the passage "Pohjantuuli ja Aurinko" ("Northwind and the Sun" in Finnish).

Statistical analysis

IBM SPSS Statistics 29×64 was used to conduct the statistical analysis. The Wilcoxon signed rank test was employed to analyze the changes in AVQI and ABI scores between the baseline and follow-up periods. The Wilcoxon signed rank test was also employed to compare subparameters of AVQI and ABI between the baseline and follow-up periods. Spearman's correlation test was used to examine the relationship between the change in AVQI and ABI scores and the change in the Hoehn and Yahr scores between the two measurement points. Similarly, Spearman's correlation test was used to examine the relationship between the change in AVQI and ABI scores and the change in UPDRS-III scale scores between the two measurement points.

RESULTS

Changes in AVQI and its subparameters

The first research question asked: (1) Between the two measurement points (baseline and the approximate 1-year follow-up), does voice quality progress as measured by AVQI? Results support our hypothesis. Poorer voice quality, illustrated by increased AVQI scores, was observed in the follow-up period compared to the baseline. Baseline AVQI mean score was 1.79 (range 0.14–4.83, SD = 1.60). Follow-up AVQI mean score was 2.25 (range 0.55–4.53,

TABLE 2.
AVQI: Baseline and Follow-up Data

Speaker ID	AVQI baseline score	AVQI follow-up score
1	1.06	1.12
2	0.84	2.05
3	1.25	2.46
4	0.67	1.2
5	2.01	3.3
6	0.24	0.55
7	3.69	4.53
8	4.83	4.01
9	3.27	2.64
10	0.14	0.87
11	0.16	0.79
12	3.28	3.43
Mean	1.79	2.25
Minimum	0.14	0.55
Maximum	4.83	4.53
SD	1.60	1.36

Abbreviation: AVQI, Acoustic Voice Quality Index.

SD = 1.36). Speaker baseline and follow-up AVQI scores are reported in Table 2. Results of the Wilcoxon signed rank test indicated a significant difference in AVQI scores between the baseline and follow-up period ($Z = -2.002$, $P = 0.045$). Examination of the change in AVQI subparameters revealed statistically significant changes in only two of the six subparameters: HNR ($Z = -1.961$, $P = 0.050$) and CPPs ($Z = -2.118$, $P = 0.034$). AVQI subparameters for both baseline and follow-up are provided in Appendix A.

Changes in ABI and its subparameters

The second research question asked: Between the two measurement points (baseline and the approximate 1-year follow-up), does voice quality progress as measured by ABI? Results support our hypothesis. Increased breathiness, as indicated by increased ABI scores, was observed in the follow-up period compared to the baseline. Baseline ABI mean score was 2.92 (range 1.27–5.31, SD = 1.57). Follow-up ABI mean score was 3.42 (range 1.40–5.40, SD = 1.38). Speaker baseline and follow-up ABI scores are reported in Table 3. Results of the Wilcoxon signed rank test showed a significant difference in ABI scores between the baseline and follow-up period ($Z = -2.197$, $P = 0.028$). Examination of the change in ABI subparameters revealed statistically significant changes only in CPPs ($Z = -2.118$, $P = 0.034$). ABI subparameters for both baseline and follow-up are provided in Appendix B.

AVQI, ABI, and measures of disease stage

The third research question asked: Are there correlations between the change in AVQI and ABI scores and the change in measures of PD stage (Hoehn and Yahr and UPDRS-III) over the measurement period? Results support our hypothesis.

TABLE 3.
ABI: Baseline and Follow-up Data

Speaker ID	ABI baseline score	ABI follow-up score
1	1.73	2.25
2	2.17	3.39
3	1.58	3.20
4	2.24	2.56
5	3.97	5.34
6	1.52	1.98
7	5.31	5.40
8	5.29	4.73
9	3.19	3.79
10	1.75	2.25
11	1.27	1.40
12	4.98	4.75
Mean	2.92	3.42
Minimum	1.27	1.40
Maximum	5.31	5.40
SD	1.57	1.38

Abbreviation: ABI, Acoustic Breathiness Index.

Results of Spearman's correlation test revealed no correlation between the change in Hoehn and Yahr and AVQI ($r_s = 0.292$, $P = 0.357$) nor between Hoehn and Yahr and ABI ($r_s = 0.477$, $P = 0.117$). Additionally, no correlation was found between change in UPDRS-III and AVQI ($r_s = -0.063$, $P = 0.847$) nor between UPDRS-III and ABI ($r_s = -0.275$, $P = 0.388$). For this analysis, the change between baseline and follow-up for each analyzed score (AVQI, ABI, Hoehn and Yahr, YPDRS-III) was found prior to running Spearman's correlation test. Follow-up and baseline data for Hoehn and Yahr scores and UPDRS-III scores are reported in Table 4.

Percentage of change in speakers over cutoff score for AVQI and ABI

The AVQI cutoff score in the analysis of Finnish language speakers is 1.83. At baseline, five of the 12 speakers

TABLE 4.
PD Rating Scale Scores, Baseline and Follow-up

	Baseline		Follow-up	
Speaker ID	H & Y	UPDRS-III	H & Y	UPDRS-III
1	2.5	12	3.0	14
2	0.0	6	1.0	8
3	1.5	16	2.5	21
4	2.0	17	2.5	20
5	1.0	5	1.5	7
6	2.0	8	3.0	15
7	2.5	19	3.0	30
8	2.5	12	3.0	16
9	2.0	7	2.5	9
10	2.0	8	2.0	8
11	1.0	5	1.5	6
12	2.0	13	2.0	17

Abbreviations: H & Y, Hoehn and Yahr; UPDRS-III, Unified Parkinson's Disease Rating Scale III.

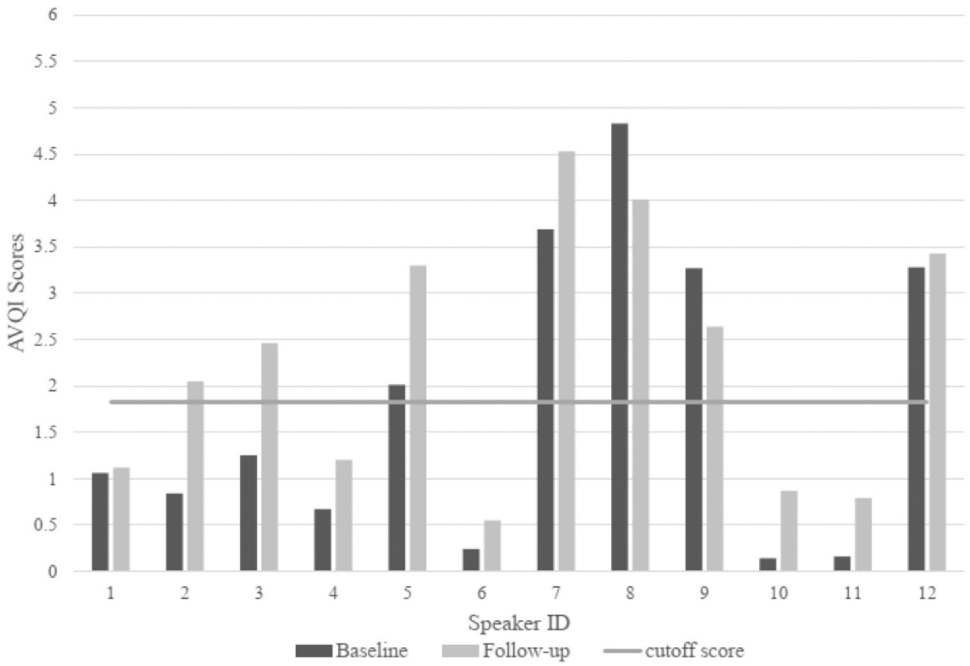

FIGURE 1. AVQI cutoff, baseline, and follow-up scores.

(41.67%) had AVQI scores greater than the cutoff score. At follow-up, seven of the 12 speakers (58.33%) had AVQI scores greater than the cutoff score. Results are illustrated in Figure 1.

The ABI cutoff score in the analysis of Finnish language speakers is 2.68. At baseline, five of the 12 speakers (41.67%) had ABI scores greater than the cutoff score. At follow-up, seven of the 12 speakers (58.33%) had ABI scores greater than the cutoff score. These results are illustrated in Figure 2.

DISCUSSION

The purpose of this pilot study was to examine voice changes in early-stage PD. Analysis was completed across two measurement points (approximately a 1-year period)

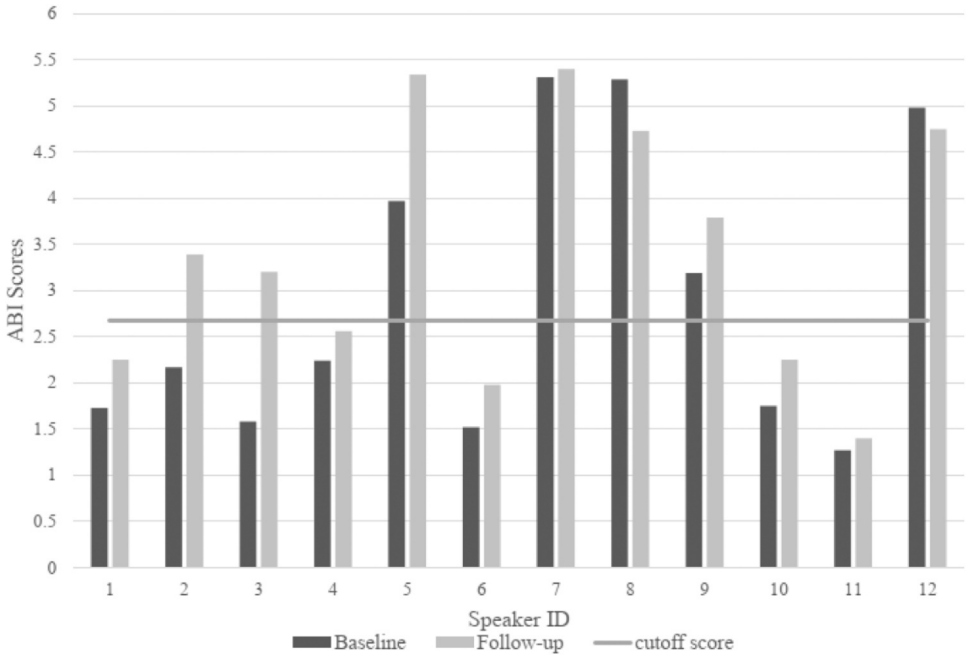

FIGURE 2. ABI cutoff, baseline, and follow-up scores.

utilizing two acoustic measures, AVQI and ABI. These measures of voice were then analyzed in comparison to measures of PD severity as measured by the Hoehn and Yahr and UPDRS-III scales. Results illustrated significant changes in both AVQI and ABI scores across the measurement period. No correlation was found at either baseline or follow-up between AVQI or ABI and the measures of PD severity.

In the first research question, we examined changes in AVQI over the course of 1-year in individuals with early-stage PD. A statistically significant change was observed in AVQI scores as well as in two of the six AVQI parameters, HNR and CPPs, between the two measurement points. The aligning trends between CPPs and AVQI are expected, given CPPs is the most influential parameter in the AVQI algorithm.[27] The increased number of speakers who had AVQI scores above the cutoff (five speakers at baseline, seven speakers at follow-up) in part illustrate the progression of voice changes in early-stage PD. In two speakers, though their AVQI scores at both measurement points were above the cutoff, their follow-up scores were lower than their baseline scores. This could possibly be the result of medication or a difference in timing between when medication was taken and when the speech samples were recorded.[10]

The second research question, examined changes in ABI over the course of 1-year in individuals with early-stage PD. In analysis of ABI scores, a statistically significant change was observed between the two measurement points. CPPs was the only ABI parameter that showed significant change between the two measurement points. CPPs and CPP have previously been proven to be strong predictors of breathiness and dysphonia.[28] This result encourages reflection on the application of multiparametric measures, like ABI in the clinical setting, or if single parameters, like CPP and CPPs, are sufficient in analyzing and tracking changes in voice. Additionally, like AVQI results, two of the speakers presented with lower ABI scores at follow-up compared to baseline, though both scores were above the cutoff.

Changes in HNR and CPPs were found to be statistically significant between the two points of measurement. HNR is a measure of perturbation and perceived as hoarseness and hypophony, weak voice.[11,12] This result is supported by the findings in Rusz (2011), which found HNR to be one of several measures of perturbation which distinguished healthy controls from individuals with early, untreated PD. The change observed in HNR could indicate changes in airflow and vocal fold control in early-stage PD. CPPs examines the severity of dysphonia, specifically the quality of breathiness, in voice.[28,29] This result indicates the progression of breathiness in the early-stage PD speakers. Though HNR and CPPs are parameters within AVQI and ABI, their value as single parameters in analysis of early-stage PD voice should be further examined.

Changes in voice have been observed up to 5 years prior to the development of motor symptoms.[1] In the current study, 10 of the speakers were diagnosed with PD when the baseline data was collected. The remaining two speakers received a formal PD diagnosis within 1 month of baseline data being collected. At baseline five speakers (41.67%) had an AVQI and ABI score greater than the cutoff score, indicating dysphonic and breathy voice in these speakers. Given that PD remains a clinical diagnosis, reliant on observable motor and nonmotor symptoms, it is not surprising that a portion of speakers presented with a degree of dysphonia at baseline.

In a study by Rusz and colleagues (2011), speech and voice changes were examined in early-stage untreated PD speakers. Results revealed that 78% of the participating speakers presented with impairments in speech and voice. In further examination of the results, the most commonly impacted speech subsystem was prosody.[11] The causes of prosodic and voice quality changes in PD differ. Dysprosody occurs secondary to changes in pitch, decreased variation in fundamental frequency, and stress.[2,30] Changes in voice quality can arise from a multitude of changes in the vocal mechanism including phase asymmetry (rough, asthenic voice), and incomplete glottal closure (breathy voice).[2] The results from the abovementioned study and this current study show changes in voice in early-stage PD. This encourages further research on the specific communication dysfunctions that arise in early-stage PD. Examination of both speech and voice features is warranted in the clinical setting with speakers with PD. The influence of medical intervention on the early changes should also be investigated.

The relationship between PD progression and changes in speech and voice is variable, with only some parameters becoming more severe as the disease progresses.[14] In analysis of progression of dysphonia severity and disease stage, results from the current study illustrated an increase in both AVQI and ABI scores though this progression was not found to be correlated with the measures of PD severity (Hoehn and Yahr scale, UPDRS-III scale). A follow-up study conducted by Skodda and colleagues (2013), examined speech and voice changes over a period of time in PD speakers. The participating speakers had been diagnosed with PD between 1 and 20 years (mean = 6.09), and there was a range of 12 to 88 months (mean = 32.25) between data collection points. Results of the study by Skodda and colleagues (2013) illustrated that while several parameters of voice did increase in severity, including shimmer, jitter, noise-to-harmonics ratio, parameters of speech, such as decreased speech rate and deterioration of articulation, and UPDRS-III scores remained relatively

consistent. These findings further illustrate the discrepancy between voice severity and disease stage. Given that the UPDRS-III scale is completed by physician observation, it is important to consider how clinical experience affects this perceptual scoring method.

In interpretation of these results, it is important to keep in mind several limiting factors. Secondary to the ongoing nature of AAVISTUS, limited data were available at the time of the study. Given the limited number of speakers in this study, future research on a larger speaker pool is warranted to expand our understanding of how voice changes in early PD influence multiparametric acoustic measures. Another limiting factor is the age range of the speakers. The youngest speaker included in the study was diagnosed with PD at 49.58 years of age, the oldest speaker was diagnosed with PD at 84.33 years of age. In consideration of the results, it is possible that the age of the speakers was influential in the results. It is known that changes in voice occur as individuals age, including changes in parameters like fundamental frequency and shimmer.[31,32] These parameters are also known to be influenced by PD.[1,6,11,7] In future studies analyzing communication changes in early-stage PD, it would be beneficial to concentrate on speakers within a narrower age range. An additional limiting factor to consider is the fact that speech samples for the baseline and follow-up measurement periods were collected only once. Voice and speech parameters may change due to various situational factors. We cannot exclude the possibility that voice quality is influenced by the patient's medication either. In the ideal case, participants would be recorded after the same time interval from the time they last took their medication.

Supporting the development of objective acoustic measures for early PD detection is crucial. Acoustic measures provide a means for early detection of PD. Therefore, supporting the development of acoustic measures is crucial. However, for this to be realized we must expand our understanding of how the voice changes in the early stages of PD. It is in the development of these measures that one must also consider their application in the clinical setting. For clinical value, voice analysis must be comprehensive and concise. Under the current research conditions, ABI was found to be more sensitive in detecting early voice changes among the studied PD speakers compared to AVQI. The single measures of CPPs and HNR also changed significantly between the two measurement points. There is need for additional investigation on the benefit of multiparametric measures for the purpose of voice analysis in the clinical setting compared to single measures. We must also determine which acoustic measures should be included in the battery for voice analysis, and how these measures should vary depending on the voice disorder being analyzed.

The aim of this pilot study was to analyze voice changes in early PD utilizing AVQI and ABI and to analyze these changes in relation to Hoehn and Yahr and UPDRS-III scores. A statistically significant change in AVQI and ABI scores was seen between the baseline and follow-up periods illustrating voice quality changes, specifically changes in breathiness, vocal fold and airflow control in early-stage PD. ABI was more sensitive to changes in voice quality compared to AVQI. Further research is warranted on the specific acoustic measures that would provide the most accurate and comprehensive analysis of voice in speakers with early-stage PD for the purposes of detecting and tracking disease progression. Future studies comparing early-stage PD speakers to healthy, age-matched controls within a larger speaker group is warranted.

Data availability

Data is not available to be shared publicly due to the nature of this research. Data will be available after the active phase of the study in 2028.

Declaration of Competing Interest

No author has personal or financial interests that affects their objectivity or serves as a potential conflict in this research.

APPENDIX A. AVQI: Baseline and Follow-up Data

Baseline data								Follow-up data						
Speaker ID	AVQI	Slope of LTAS (dB)	Tilt of LTAS	Shimmer local (%)	Shimmer local (dB)	HNR (dB)	CPPs	AVQI	Slope of LTAS (dB)	Tilt of LTAS	Shimmer local (%)	Shimmer local (dB)	HNR (dB)	CPPs
1	1.06	−20.84	−9.77	1.18	0.10	32.28	21.43	1.12	−14.53	−12.6	2.77	0.3	25.78	19.18
2	0.84	−21.86	−12.79	2.55	0.22	27.07	17.39	2.05	−18.22	−11.63	6.82	0.62	22.17	15.07
3	1.25	−19.63	−11.03	3.55	0.31	22.03	17.73	2.46	−25.74	−10.74	6.32	0.55	15.28	12.09
4	0.67	−18.95	−8.22	4.4	0.40	25.23	18.91	1.2	−16.62	−8.07	3.7	0.33	22.68	17.21
5	2.01	−24.3	−10.22	2.21	0.20	29.49	13.35	3.3	−21.17	−7.75	6.44	0.55	18.94	10.87
6	0.24	−14.14	−12.21	3.07	0.27	24.3	17.77	0.55	−17.21	−12.79	3.82	0.34	21.16	16.99
7	3.69	−35.34	−9.97	5.41	0.48	20.64	11.32	4.53	−25.52	−11.1	5.19	0.48	14.13	10.41
8	4.83	−28.31	−11.18	9.83	0.86	13.16	8.68	4.01	−20.65	−11.28	13.57	1.2	8.22	8.89
9	3.27	−18.41	−12.36	10.65	0.98	10.87	13.31	2.64	−18.61	−12.06	3.18	0.28	20.66	14.16
10	0.14	−15.68	−10.43	1.75	0.16	27.74	17.47	0.87	−16	−10.76	2.34	0.2	22.2	17.05
11	0.16	−14.21	−11.72	2.29	0.20	25.09	21.18	0.79	−16.49	−10.42	4.87	0.46	20.02	19.77
12	3.28	−17.37	−11.96	7.23	0.63	13.86	11.04	3.43	−14.78	−11.9	4.65	0.46	17.34	12.69
Mean	1.79	−20.75	−10.99	4.51	0.40	22.65	15.80	2.25	−18.80	−10.93	5.31	0.48	19.05	14.53
Minimum	0.14	−35.34	−12.79	1.18	0.10	10.87	8.68	0.55	−25.74	−12.79	2.34	0.2	8.22	8.89
Maximum	4.83	−14.14	−8.22	10.65	0.98	32.28	21.43	4.53	−14.53	−7.75	13.57	1.2	25.78	19.77
SD	1.60	6.16	1.32	3.16	0.284	6.81	4.14	1.36	3.78	1.59	2.99	0.26	4.72	3.58

Abbreviations: AVQI, Acoustic Voice Quality Index; CPPs, smoothed cepstral peak prominence; HNR, harmonics-to-noise ratio; slope of LTAS, slope of long-term average spectrum; tilt of LTAS, tilt of long-term average spectrum.

APPENDIX B. ABI: Baseline and Follow-up Data

Baseline data											Follow-up data									
Speaker ID	ABI	H1-H2	Jitt local	Shimmer %	Shimmer dB	PSD	HNR-D	GNE	Hfno	CPPs	ABI	H1-H2	Jitt local	Shimmer %	Shimmer dB	PSD	HNR-D	GNE	Hfno	CPPs
1	1.73	6.97	0.25	1.18	0.10	0.06	37.57	0.95	2.21	21.43	2.25	6.41	0.35	2.77	0.30	0.08	38.07	0.97	2.49	19.18
2	2.17	3.69	0.25	2.55	0.22	0.05	37.33	0.88	2.81	17.39	3.39	3	0.49	6.82	0.62	0.06	37.79	0.84	2.45	15.07
3	1.58	−1.01	0.49	3.55	0.31	0.19	22.69	0.95	2.24	17.73	3.20	4.32	3.72	6.32	0.55	0.35	22.51	0.89	2.28	12.09
4	2.24	1.94	0.34	4.40	0.40	0.08	36.46	0.94	1.89	18.91	2.56	−1.62	0.24	3.70	0.33	0.11	34.02	0.9	1.95	17.21
5	3.97	19.85	0.33	2.21	0.20	0.06	34.49	0.78	2.38	13.35	5.34	13.5	0.55	6.44	0.55	0.10	30.15	0.53	1.81	10.87
6	1.52	3.81	0.39	3.07	0.27	0.07	37.16	0.97	2.49	17.77	1.98	5.59	0.33	3.82	0.34	0.10	36.13	0.97	2.65	16.99
7	5.31	13.45	0.65	5.41	0.48	0.24	22.03	0.59	2.24	11.32	5.40	5.83	0.93	5.19	0.48	0.27	21.06	0.74	2.14	10.41
8	5.29	10.78	1.22	9.83	0.86	0.35	16.91	0.58	2.33	8.68	4.73	12.25	1.93	13.57	1.20	0.27	17.93	0.51	2.29	8.89
9	3.19	−8.15	1.94	10.65	0.98	0.41	21.26	0.91	2	13.31	3.79	6.16	0.66	3.18	0.28	0.14	31.22	0.89	2.19	14.16
10	1.75	3.46	0.26	1.75	0.16	0.11	40.2	0.97	2.15	17.47	2.25	1.35	0.30	2.34	0.20	0.05	34.24	0.96	2.40	17.05
11	1.27	5.22	0.28	2.29	0.20	0.1	33.22	0.98	2.55	21.18	1.40	6.48	1.11	4.87	0.46	0.17	28.69	0.98	2	19.77
12	4.98	−2.38	0.58	7.23	0.63	0.05	31.79	0.50	2.01	11.04	4.75	0.99	0.46	4.65	0.46	0.08	37.28	0.72	2.25	12.69
Mean	2.92	4.80	0.58	4.51	0.40	0.15	30.93	0.83	2.28	15.80	3.42	5.36	0.92	5.31	0.48	0.15	30.76	0.83	2.24	14.53
Minimum	1.27	−8.15	0.25	1.18	0.10	0.05	16.91	0.50	1.89	8.68	1.40	−1.62	0.24	2.34	0.20	0.05	17.93	0.51	1.81	8.89
Maximum	5.31	19.85	1.94	10.65	0.98	0.41	40.2	0.98	2.81	21.43	5.40	13.50	3.72	13.57	1.20	0.35	38.07	0.98	2.65	19.77
SD	1.57	7.43	0.51	3.16	0.28	0.12	7.96	0.18	0.26	4.14	1.38	4.34	1.00	2.99	0.26	0.10	6.93	0.17	0.24	3.58

Abbreviations: ABI, Acoustic Breathiness Index; CPPs, smoothed cepstral peak prominence; GNE, glottal-to-noise excitation ratio; Hfno, high-frequency noise of 6000 Hz; HNR-D, harmonics-to-noise ratio of Dejonckere; H1-H2, amplitude difference between the first two harmonics; Jitt local, jitter local; PSD, period standard deviation.

References

1. Harel B, Cannizzaro M, Snyder PJ. Variability in fundamental frequency during speech in prodromal and incipient Parkinson's disease: a longitudinal case study. *Brain Cogn.* 2004;56:24–29. https://doi.org/10.1016/j.bandc.2004.05.002.
2. Ma A, Lau KK, Thyagarajan D. Voice changes in Parkinson's disease: what are they telling us? *J Clin Neurosci.* 2020;72:1–7. https://doi.org/10.1016/j.jocn.2019.12.029.
3. Logemann JA, Fisher HB, Boshes B, et al. Frequency and cooccurrence of vocal tract dysfunctions in the speech of a large sample of Parkinson patients. *J Speech Hear Disord.* 1978;43:47–57. https://doi.org/10.1044/jshd.4301.47.
4. Schalling E, Johansson K, Hartelius L. Speech and communication changes reported by people with Parkinson's disease. *Folia Phoniatr Logop.* 2017;69:131–141. https://doi.org/10.1159/000479927.
5. Fox CM, Ramig LO. Vocal sound pressure level and self-perception of speech and voice in men and women with idiopathic Parkinson disease. *Am J Speech Lang Pathol.* 1997;6:85–94. https://doi.org/10.1044/1058-0360.0602.85.
6. Ramig L, Halpern A, Spielman J, et al. Speech treatment in Parkinson's disease: randomized controlled trial (RCT). *Mov Disord.* 2018;33:1777–1791. https://doi.org/10.1002/mds.27460.
7. Skodda S, Grönheit W, Mancinelli N, et al. Progression of voice and speech impairment in the course of Parkinson's disease: a longitudinal study. *Parkinson's Dis.* 2013;2013:389195. https://doi.org/10.1155/2013/389195.
8. Patel RR, Sundberg J, Gill B, et al. Glottal airflow and glottal area waveform characteristics of flow phonation in untrained vocally healthy adults. *J Voice.* 2022;36:140.e1–140.e21. https://doi.org/10.1016/j.jvoice.2020.07.037.
9. Ho AK, Iansek R, Marigliani C, et al. Speech impairment in a large sample of patients with Parkinson's disease. *Behav Neurol.* 1999;11:131–137. https://doi.org/10.1155/1999/327643.
10. Tracy JM, Özkanca Y, Atkins DC, et al. Investigating voice as a biomarker: deep phenotyping methods for early detection of Parkinson's disease. *J Biomed Inform.* 2020;104:103362. https://doi.org/10.1016/j.jbi.2019.103362.
11. Rusz J, Cmejla R, Ruzickova H, et al. Quantitative acoustic measurements for characterization of speech and voice disorders in early untreated Parkinson's disease. *J Acoust Soc Am.* 2011;129:350–367. https://doi.org/10.1121/1.3514381.
12. Rusz J, Tykalová T, Novotný M, et al. Distinct patterns of speech disorder in early-onset and late-onset de-novo Parkinson's disease. *Npj Parkinson's Dis.* 2021;7:1. https://doi.org/10.1038/s41531-021-00243-1.
13. Hemmerling D, Wojcik-Pedziwiatr M. Prediction and estimation of Parkinson's disease severity based on voice signal. *J Voice.* 2022;36:439.e9–439.e20. https://doi.org/10.1016/j.jvoice.2020.06.004.
14. Holmes RJ, Oates JM, Phyland DJ, et al. Voice characteristics in the progression of Parkinson's disease. *Int J Lang Commun Disord.* 2000;35:407–418. https://doi.org/10.1080/136828200410654.
15. Kankare E, Laukkanen A-M. Validation of the Acoustic Breathiness Index in speakers of Finnish language. *J Clin Med.* 2023;12:24. https://doi.org/10.3390/jcm12247607.
16. Kankare E, Rantala L, Ikävalko T, et al. Akustisen äänenlaatuindeksin (AVQI) version 03.01 validointi suomenkielisille puhujille. *Puhe ja kieli.* 2020;3:3. https://doi.org/10.23997/pk.101516.
17. Maryn Y, Corthals P, Van Cauwenberge P, et al. Toward improved ecological validity in the acoustic measurement of overall voice quality: combining continuous speech and sustained vowels. *J Voice.* 2010;24:540–555. https://doi.org/10.1016/j.jvoice.2008.12.014.
18. Maryn Y, De Bodt M, Roy N. The Acoustic Voice Quality Index: toward improved treatment outcomes assessment in voice disorders. *J Commun Disord.* 2010;43:161–174. https://doi.org/10.1016/j.jcomdis.2009.12.004.
19. Boersma P, Weenink D. *PRAAT [Computer Program]: Doing Phonetics by Computer (Version 6.0.37)*. Amsterdam, The Netherlands: Institute of Phonetic Sciences; 2018.⟨http://www.praat.org⟩.
20. Mayer, J. (2023). VOXplot, Acoustic Voice Quality Profile (v2.0.1 (2023-09)) [Computer software]. © lingphon/Jörg Mayer. Available at: https://voxplot.lingphon.com/en/. Accessed December 5, 2023.
21. Boutsen F, Park E, Dvorak JD. An efficacy study of voice quality using cepstral analyses of phonation in Parkinson's disease before and after SPEAK-OUT!®. *Folia Phoniatr Logop Off Organ Int Assoc Logop Phoniatr (IALP).* 2023;75:35–42. https://doi.org/10.1159/000525884.
22. Moya-Galé G, Spielman J, Ramig LA, et al. The Acoustic Voice Quality Index (AVQI) in people with Parkinson's disease before and after intensive voice and articulation therapies: secondary outcome of a randomized controlled trial. *J Voice.* 2022. https://doi.org/10.1016/j.jvoice.2022.03.014.ISSN 0892-1997.
23. Convey RB, Laukkanen A-M, Ylinen S, et al. Analysis of voice in Parkinson's disease utilizing the Acoustic Voice Quality Index. *J Voice.* 2024. https://doi.org/10.1016/j.jvoice.2023.12.025.ISSN 0892-1997.
24. Barsties v. Latoszek B, Maryn Y, Gerrits E, et al. The Acoustic Breathiness Index (ABI): a multivariate acoustic model for breathiness. *J Voice.* 2017;31:511.e11–511.e27. https://doi.org/10.1016/j.jvoice.2016.11.017.
25. Anttila, M., Hartikainen, S., Herukka, S.-K., et al. (2020). Parkinsonin taudin lääkkeiden erityiskorvausoikeuden saajien määrän muutos vuosina 1996–2015. Available at: https://erepo.uef.fi/handle/123456789/23656. Accessed May 17, 2024.
26. Boersma, P., & Weenink, D. (2022). Praat: doing phonetics by computer [Computer program] (6.2.05) [Computer software]. Available at: http://www.praat.org/. Accessed March 7, 2024.
27. Maryn Y, Weenink D. Objective dysphonia measures in the program Praat: smoothed cepstral peak prominence and Acoustic Voice Quality Index. *J Voice.* 2015;29:35–43. https://doi.org/10.1016/j.jvoice.2014.06.015.
28. Barsties v Latoszek B, Maryn Y, Gerrits E, et al. A meta-analysis: acoustic measurement of roughness and breathiness. *J Speech Lang Hear Res.* 2018;61:298–323. https://doi.org/10.1044/2017_JSLHR-S-16-0188.
29. Šimek M, Rusz J. Validation of cepstral peak prominence in assessing early voice changes of Parkinson's disease: effect of speaking task and ambient noise. *J Acoust Soc Am.* 2021;150:4522–4533. https://doi.org/10.1121/10.0009063.
30. Galaz Z, Mekyska J, Mzourek Z, et al. Prosodic analysis of neutral, stress-modified and rhymed speech in patients with Parkinson's disease. *Comput Methods Programs Biomed.* 2016;127:301–317. https://doi.org/10.1016/j.cmpb.2015.12.011.
31. Rojas S, Kefalianos E, Vogel A. How does our voice change as we age? A systematic review and meta-analysis of acoustic and perceptual voice data from healthy adults over 50 years of age. *J Speech Lang Hear Res.* 2020;63:533–551. https://doi.org/10.1044/2019_JSLHR-19-00099.
32. Verdonck-de Leeuw IM, Mahieu HF. Vocal aging and the impact on daily life: a longitudinal study. *J Voice Off J Voice Found.* 2004;18:193–202. https://doi.org/10.1016/j.jvoice.2003.10.002.