Article

Multivariate Analysis of Vocal Fold Vibrations on Various Voice Disorders Using High-Speed Digital Imaging

Akihito Yamauchi [1,*], Hiroshi Imagawa [1], Hisayuki Yokonishi [2], Ken-Ichi Sakakibara [3] and Niro Tayama [4]

[1] The Department of Otorhinolaryngology, The University of Tokyo Hospital, Tokyo 113-8655, Japan; imagawa@m.u-tokyo.ac.jp

[2] The Department of Otolaryngology, Tokyo Metropolitan Bokutoh Hospital, Tokyo 130-8575, Japan; h_yokonishi@yahoo.co.jp

[3] The Department of Communication Disorders, Health Sciences University of Hokkaido, Hokkaido 061-0293, Japan; quesokis@gmail.com

[4] The Department of Otolaryngology and Tracheo-Esophagology, National Center for Global Health and Medicine, Tokyo 162-8655, Japan; ntayama@hosp.ncgm.go.jp

* Correspondence: drachilles23@yahoo.co.jp; Tel.: +81-3-3815-5411

Abstract: Although many quantitative parameters have been devised to describe abnormalities in vocal fold vibration, little is known about the priority of these parameters. We conducted a prospective study using high-speed digital imaging to elucidate disease-specific key parameters (KPs) to characterize the vocal fold vibrations of individual voice disorders. From 304 patients with various voice disorders and 46 normal speakers, high-speed digital imaging of a sustained phonation at a comfortable pitch and loudness was recorded and parameters from visual-perceptual rating, laryngotopography, digital kymography, and glottal area waveform were calculated. Multivariate analysis was then applied to these parameters to elucidate the KPs to explain each voice disorder in comparison to normal subjects. Four key parameters were statistically significant for all laryngeal diseases. However, the coefficient of determination (R^2) was very low (0.29). Vocal fold paralysis (8 KPs, $R^2 = 0.76$), sulcus vocalis (4 KPs, $R^2 = 0.74$), vocal fold scarring (1 KP, $R^2 = 0.68$), vocal fold atrophy (6 KPs, $R^2 = 0.53$), and laryngeal cancer (1 KP, $R^2 = 0.52$) showed moderate-to-high R^2 values. The results identified different KPs for each voice disorder; thus, disease-specific analysis is a reasonable approach.

Keywords: vocal fold vibration; voice disorder; high-speed digital imaging; digital kymography; laryngotopography

1. Introduction

A person's voice results from the interaction and matching of respiration, vocal fold vibrations, and speech movements. Vocal fold vibrations are the main generator of glottal sound; thus, the evaluation of these vibrations plays a pivotal role in the detection, diagnosis, and treatment of voice disorders.

Videostroboscopy (VS) is widely used in daily clinical practice to observe the rapid dynamics of vocal fold vibrations [1]. However, the low sampling rate leads to low temporal data quality and limits its assessment based on a subjective visual perceptual rating (VPR). Furthermore, VS is limited to specific patient populations because it is inapplicable to non-sustained phonations or moderate-to-severe dysphonia [2,3].

Another common imaging approach is high-speed digital imaging (HSDI). HSDI registers actual vocal fold vibrations with a high sampling rate (2000–10,000 Hz). It provides rich temporal information and permits reliable quantitative analysis. Furthermore, since data acquisition is not dependent on the examinee's fundamental frequency, HSDI can be applied to non-stationary phonations to which VS cannot be applied [4,5]. Because of its potential in multifaceted, objective, and quantitative analysis, HSDI is considered to be the future gold standard for vocal fold vibratory assessment.

However, it is not yet widely used in clinical practice due to its unsolved disadvantages. First, a standardized analysis method has yet to be established, although various analysis methods have been proposed. Additionally, the choice of parameter is at the discretion of the researchers. Although methods and parameters should be modified depending on the research questions and hypotheses of the study, comparisons of different studies are difficult without the standardized analysis methods or parameters that should be at the core of HSDI research. Various quantitative parameters have been proposed, and too many existing parameters make it difficult to choose which to measure as well as obscuring key parameters (parameters of primary importance) [4,5]. To balance the merits and demerits of HSDI, it is essential to extract the most important parameters. There is a relative scarcity of studies proving the merits of HSDI and proving the superiority of HSDI to VS. The high cost and time-consuming nature of data processing and analysis of HSDI are other disadvantages that come with this approach.

Bohr et al. performed a set of HSDI studies to address the matter of standardization (see [6,7]). They tested 20 parameters from the glottal area waveform (GAW) and 16 parameters from phonovibrography (PVG) in 207 subjects (55 males, 152 females) using univariate analyses. They selected five candidate GAW parameters and eight PVG parameters from male subjects and seven GAW parameters and 13 PVG parameters from female subjects to distinguish normal and abnormal vocal fold vibrations. However, these studies did not prioritize the parameters by multivariate analysis; thus, the key parameters for each disorder are unclear. Other popular analysis methods, such as VPR and digital kymography (DKG), were not utilized in these studies.

In another attempt to establish the standardization of analysis methods and parameters, the authors investigated the vibratory characteristics of various voice disorders compared to those in normal subjects using VPR, DKG, laryngotopography (LTG), and GAW in previous studies (see [8–12]). Through these study series, we developed a parameter set for HSDI derived from these four analysis methods. Further, using the parameter set, we clarified the disease-specific vibratory characteristics of various laryngeal disorders in comparison with normal subjects by univariate analyses. However, the prioritization of vibratory parameters using multivariate analyses has not been attempted.

In this context, we investigated the following hypotheses: First, HSDI is superior to VS in the assessment of abnormal vocal fold vibrations. Although this superiority has been proven in the past, only a few studies have been carried out to support this. This hypothesis can be justified when the assessability of vocal fold vibrations of HSDI is significantly higher than that of VS. Second, HSDI-derived parameters can distinguish between normal and abnormal vocal fold vibrations. This hypothesis can be proven when there are HSDI parameters with statistically significant differences between the normal and abnormal groups. Third, HSDI can detect vibratory characteristics that are specific to each voice disorder. This hypothesis signifies that different sets of key HSDI parameters are found for each voice disorder. Fourth, an analysis method using multiple techniques is more effective than analysis using a single method. This hypothesis can be proven if the key parameters are derived using multiple analysis methods.

To prove these hypotheses, we performed a comparative study of normal subjects and patients with various voice disorders, using HSDI to extract parameters to describe the disease-specific vibratory abnormalities of each voice disorder by multivariate analysis. Further, we compared the assessability of vocal fold vibrations between VS and HSDI in the patient group to demonstrate the superiority of HSDI to VS.

2. Materials and Methods

2.1. Subjects

Data were collected during routine clinical practice at the vocal outpatient clinic of the Department of Otolaryngology and Head and Neck Surgery at the University of Tokyo Hospital (Tokyo, Japan) between 2006 and 2013. During the same period, healthy volunteers without any vocal complaints and with no history of laryngeal disorders or

laryngeal pathology on video-laryngoscopy during the same period were recruited as the normal control group.

In the present study, 304 patients with voice disorders and 46 normal subjects were examined. The breakdown of disorders in these 304 patients is summarized in Table 1.

Table 1. Subject number by voice disorder.

Voice Disorder	N	Gender (Male, Female)	Age (Mean $\pm$ Standard Deviation)
Normal	46	17, 29	47.1 ± 23.5
Abnormal Total	304	198, 106	61.0 ± 15.5
Vocal fold paralysis	106	68, 38	61.0 ± 14.8
Vocal fold atrophy	61	46, 15	65.0 ± 14.9
Sulcus vocalis	29	23, 6	68.7 ± 12.3
Vocal fold polyp	20	13, 7	54.6 ± 15.7
Laryngeal cancer	15	15, 0	69.7 ± 10.1
Laryngeal leukoplakia	13	13, 0	68.0 ± 11.6
Vocal fold scar	12	2, 10	52.4 ± 13.7
Other diseases	48	18, 30	51.0 ± 15.5

2.2. High-Speed Digital Imaging

A high-speed digital camera (Photron, FASTCAM-1024PCI, Tokyo, Japan) was connected to a rigid endoscope (#4450.501, Richard Wolf, Baden-Württemberg, Germany) via an attachment lens (f = 35 mm, Nagashima Medical Instrument Corporation, Tokyo, Japan). Recordings were performed under illumination with a 300-W xenon light source at a frame rate of 4500 fps, with a spatial resolution of 512 $\times$ 400 pixels, 8-bit grayscale, and a maximum recording duration of 1.86 s.

HSDI data were recorded trans-orally during sustained phonation of the vowel at a comfortable pitch and loudness. A segment of stable vocal fold vibrations with good glottal exposure was selected, and the image sequences were analyzed using a MATLAB-based analysis program equipped with VPR, DKG, LTG, and GAW (Figure 1).

2.3. Parameters

Vocal fold vibration has been assessed by focusing on periodicity, glottal closure, closed phase, glottal gap, amplitude, mucosal wave, phase difference (PD), and other anatomical findings (supraglottic hyperactivity, secretion, atrophic change, edematous change, and non-vibrating area) [3]. The parameter set employed in this study was mostly based on the traditional assessment of vocal fold vibration. However, the set was updated and modified to suit HSDI analysis based on a literature review of the relevant field and the authors' clinical experience with voice disorders (Tables 2–6; see [13–22]).

Parameters related to size (e.g., amplitude) were normalized by the vocal fold length; these size-parameter names were labeled with the prefix N_L-. Vocal fold length was defined as the distance between the vocal process and the anterior commissure. Parameters related to time (e.g., the opening phase) were normalized by the glottal cycle; these time-parameter names were labeled with the prefix N_G-. Parameters related to both size and time (e.g., integral glottal width) were normalized by the vocal fold length and glottal cycle; these size-and-time parameter names were labeled with the prefix, N_{GL}-.

Categorical parameters were judged subjectively by examiners, and user inputs were required. The analysis was performed by two phono-surgeons who were familiar with the analysis of HSDI data. Then, the values of the two examiners were averaged. The order of the analysis was randomized, and the examiners were blinded to the subjects' information. The inter-rater reliability was calculated. Ten percent of the data were reexamined to calculate the intra-rater reliability.

Continuous parameters were automatically calculated via a MATLAB-based program, although some manual inputs were needed (e.g., to omit noise by secretion or adjust the length or area to measure).
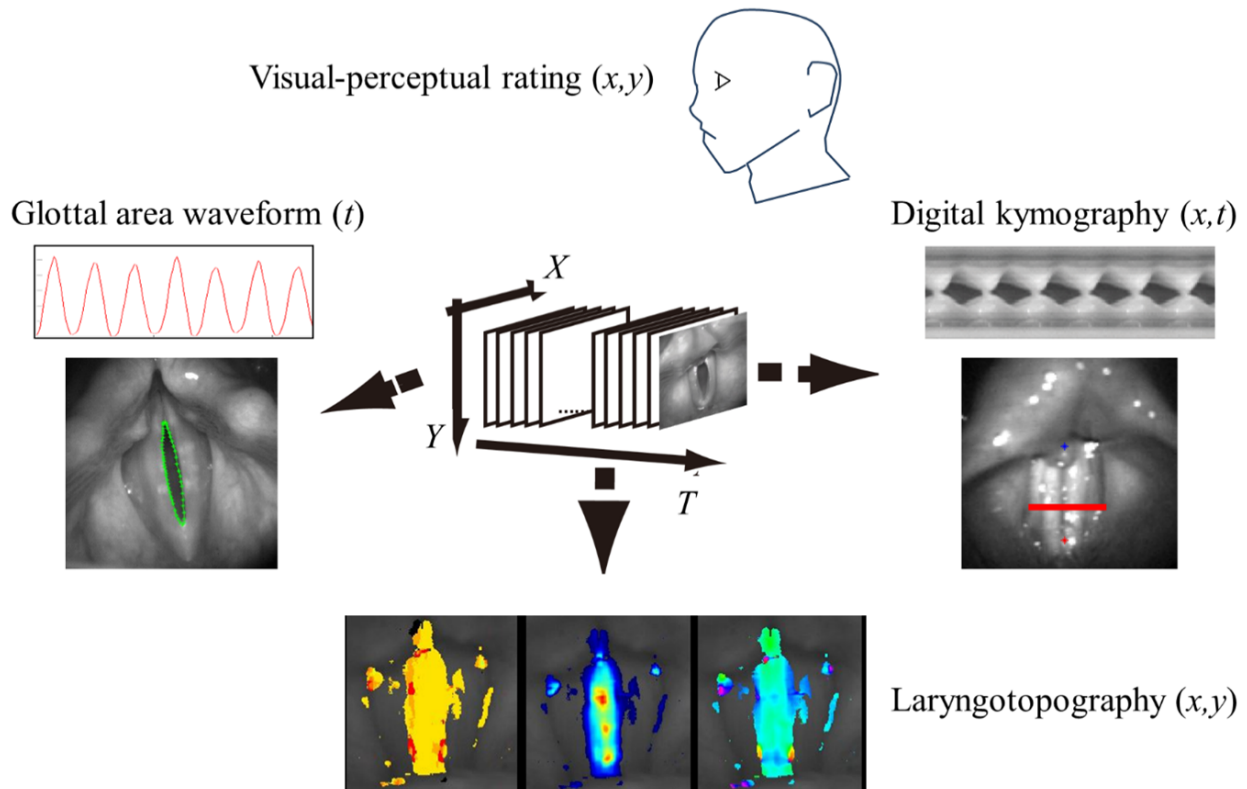

Figure 1. Analysis Method. HSDI data are composed of mediolateral (*x*-axis), longitudinal (*y*-axis), and temporal (*t*-axis) dimensions. Visual-perceptual rating is easy and effective in assessing *x*- and *y*-axis data although it is subjective and not apt for temporal analysis. Digital kymography is a classical, popular method that is effective in the quantitative assessment of *x*- and *t*-axis data. Laryngotopography is an intuitive easy-to-understand method that enables spatial assessment (*x*- and *y*-axis data). Glottal area waveform is another classical popular method that allows quantitative evaluation of temporal change of glottal area. Combining these methods guarantees thorough analysis of HSDI data.

2.4. Visual-Perceptual Rating

The VPR is the most fundamental method for HSDI data analysis, in which the vibratory characteristics are rated subjectively by examiners' close inspection of slow-motion or frame-by-frame playback of HSDI videos. This method is considered to be effective in the analysis of *x*-axis and *y*-axis data but not of *t*-axis data [13].

In this study, VPR was performed using an assessment form devised by the authors' voice team [13]. This assessment form contained 21 parameters (Table 2): periodicity, glottal closure, closed phase, glottal gap (size and location), amplitude (mean and left-right difference), mucosal wave (mean and left-right difference), lateral PD (magnitude and direction), longitudinal PD (mean, direction, and left-right difference), axis shift (magnitude and direction), and other findings (supraglottic hyperactivity, secretion, atrophic change, edematous change, and non-vibrating area). The VPR-derived parameters were labeled with the suffix VPR (e.g., glottal closure VPR).

The analysis was performed by two phono-surgeons who were familiar with the analysis of HSDI data. The order of the analysis was randomized and the examiners were blinded to the subjects' information.

Table 2. Parameters from visual-perceptual rating (VPR).

Parameter from VPR (N = 21)	Details
Categorical variables	
Periodicity [VPR]	0: periodic, 1: not periodic
Glottal closure type [VPR]	0: complete, 1: incomplete
Closed period [VPR]	0: none, 1: shortened, 2: moderate, 3: prolonged
Glottal gap size [VPR]	0: none, 1: small, 2: middle, 3: large
Glottal gap location [VPR]	0: anterior, 1: mid-glottal, 2: posterior, 3: hourglass, 4: irregular
Amplitude mean [VPR]	0: none, 1: small, 2: moderate, 3: large
Amplitude difference [VPR]	0: none, 1: small, 2: moderate, 3: large
Mucosal wave mean [VPR]	0: none, 1: small, 2: moderate, 3: large
Mucosal wave difference [VPR]	0: none, 1: small, 2: moderate, 3: large
Lateral PD magnitude [VPR]	0: none, 1: small, 2: moderate, 3: large
Lateral PD direction [VPR]	0: left-to-right, 1: right-to-left
Longitudinal PD mean [VPR] (*1)	0: none, 1: small, 2: moderate, 3: large
Longitudinal PD direction [VPR]	0: posterior-to-anterior, 1: anterior-to-posterior
Longitudinal PD difference [VPR]	0: none, 1: small, 2: moderate, 3: large
Axis shift magnitude [VPR]	0: none, 1: small, 2: moderate, 3: large
Axis shift direction [VPR]	0: left-to-right, 1: right-to-left
Supraglottal hyperactivity [VPR]	0: none, 1: mild, 2: moderate, 3: severe
Glottal secretion [VPR]	0: none, 1: small, 2: moderate, 3: large
Non-vibrating area [VPR]	0: absent, 1: present
Atrophic change [VPR]	0: absent, 1: present
Edematous change [VPR]	0: absent, 1: present

[VPR] = derived from visual perceptual rating; PD = phase difference. VPR parameters are judged subjectively by examiners, and thus user inputs are needed in the entire assessment. All the VPR parameters are categorical variables. (*1) The range of longitudinal PD mean [VPR] was from −3 to 3 (posterior-to-anterior PD was the positive value and anterior-to-posterior PS was the negative value).

2.5. Laryngotopography

LTG employs fast Fourier transformation of a brightness-versus-time curve for each pixel across images to achieve a quantitative evaluation of the spatial characteristics of amplitude, frequency, and phase [14,15]. LTG is effective for objective, quantitative, *x*-axis, and *y*-axis evaluations [14,15].

Table 3. Parameters from laryngotopography (LTG).

Parameter from LTG (N = 10)	Details
Categorical variables	
Periodicity [LTG]	0: periodic, 1: quasi-periodic, 2: not periodic
Lateral PD direction [LTG]	0: left-to-right, 1: right-to-left
Long. PD direction [LTG]	0: posterior-to-anterior, 1: anterior-to-posterior
Non-vibrating area [LTG]	0: absent, 1: present
Continuous variables	
F_0 [LTG]	Hz
N_G-lateral PD magnitude [LTG]	%, range: from 0 to 100
N_G-long. PD mean [LTG] (*1)	%, range: from −100 to 100
N_G-long. PD difference [LTG]	%, range: from 0 to 100
N_G-MWP mean [LTG] (*2)	%, range: from 0 to 100
N_G-MWP difference [LTG]	%, range: from 0 to 100

F_0 = fundamental frequency; [LTG] = derived from laryngotopography; PD = phase difference; N_G- = normalized by glottal cycle; MWP = mucosal wave persistence. LTG categorical parameters are judged subjectively by examiners, and user inputs are needed. LTG continuous parameters are automatically calculated although some manual inputs are needed (e.g., to omit noise by secretion, adjust length or area to measure). (*1) In N_G-longitudinal PD mean [LTG] (%), posterior-to-anterior PD was calculated as the positive value and anterior-to-posterior PD was calculated as the negative value. (*2) N_G-MWP signifies a temporal duration of visible mucosal wave during one glottal cycle.

This study evaluated 10 parameters (Table 3): F0, periodicity, lateral PD (magnitude and direction), longitudinal PD (mean, direction, and left-right difference), mucosal wave

persistence (temporal duration of visible mucosal wave; mean and left-right difference), and non-vibrating area. These LTG-derived parameters were labeled with the suffix ^{LTG}, (e.g., F_0^{LTG}).

Table 4. Parameters from digital kymography (DKG) 1: Single-line kymography (SLK).

Parameter from SLK (N = 26)	Details
Categorical variables	
N_G-lateral PD direction ^{SLK}	0: left-to-right, 1: right-to-left
N_L-axis shift direction ^{SLK}	0: left-to-right, 1: right-to-left
Continuous variables	
F_0 ^{SLK}	Hz
SI mean ^{SLK}	The mean of lt. and rt. SIs, range: from −1 to 1
SI difference ^{SLK}	The difference between lt. and rt. SIs (absolute value)
OQ ^{SLK}	Range: from 0 to 1, OQ at the mid-glottal level
N_G-lateral PD magnitude ^{SLK}	Lateral PD/GC
N_G-MWP mean ^{SLK} (*1)	(lt. MWP + rt. MWP)/2×GC
N_G-MWP difference ^{SLK}	(lt. MWP − rt. MWP)/GC
N_G-effective MWP mean ^{SLK}	((lt. MWP − lt. opening phase) + (rt. MWP − rt. opening phase))/2×GC
N_G-effective MWP difference ^{SLK}	((lt. MWP − lt. opening phase) − (rt. MWP − rt. opening phase))/GC
N_L-Amp. mean ^{SLK}	(lt. Amp. + rt. Amp.)/2×VFL
N_L-Amp. difference ^{SLK}	(lt. Amp. − rt. Amp.)/VFL
Amplitude asymmetry ^{SLK}	(lt. Amp. − rt. Amp.)/(lt. Amp. + rt. Amp.)
N_L-maximal glottal width ^{SLK}	Maximal glottal width/VFL
N_L-MWM mean ^{SLK} (*2)	(lt. MWM + rt. MWM)/2×VFL
N_L-MWM difference ^{SLK}	(lt. MWM − rt. MWM)/VFL
N_L-effective MWM mean ^{SLK}	(lt. MWM − lt. Amp. + rt. MWM − rt. Amp.)/2×VFL
N_L-effective MWM difference ^{SLK}	((lt. MWM − lt. Amp.) − (rt. MWM − rt. Amp.))/VFL
N_L-axis shift magnitude ^{SLK} (*3)	Axis shift/VFL
Axis shift	Axis shift/(lt. Amp. + rt. Amp.)
N_{GL}-IGW ^{SLK} (*4)	IGW/VFL×GC
N_{GL}-IGW difference ^{SLK}	(lt.-half IGW − rt.-half IGW)/VFL×GC
Asymmetry index ^{SLK}	(lt.-half IGW − rt.-half IGW)/IGW
N_{GL}-lateral peak mean ^{SLK} (*5)	(lt. Amp./lt. opening phase + (lt. Amp. − axis shift)/lt. closing phase + (rt. Amp. + axis shift)/rt. opening phase + rt. Amp./rt. closing phase)/2×VFL×GC
N_{GL}-lateral peak difference ^{SLK}	(lt. Amp./lt. opening phase + (lt. Amp. − axis shift)/lt. closing phase − (rt. Amp. + axis shift)/rt. opening phase − rt. Amp./rt. closing phase)/VFL×GC

F_0 = fundamental frequency; ^{SLK} = derived from single-line kymography; SI = speed index; lt. = left; rt. = right, OQ = open quotient; N_G- = normalized by glottal cycle; PD = phase difference; GC = glottal cycle; MWP = mucosal wave persistence; N_L- = normalized by vocal fold length; Amp. = amplitude; VFL = vocal fold length; MWM = mucosal wave magnitude; N_{GL}- = normalized by glottal cycle and vocal fold length; IGW = integral glottal width. SLK categorical parameters are judged subjectively by examiners, and user inputs are needed. SLK continuous parameters are automatically calculated although some manual inputs are needed (e.g., to omit noise by secretion, adjust length or area to measure). (*1) MWP signifies a temporal duration of visible mucosal wave (a temporal parameter of mucosal wave). (*2) MWM signifies a distance of visible mucosal wave in the lateral direction (a size parameter of mucosal wave). (*3) Axis shift is a lateral shift of glottal closure between glottal closure and glottal opening (a size measure regarding asymmetry). (*4) IGW is an integral value of glottal width (a lateral distance between left and right vocal fold free edges or lateral size of glottal area) in one glottal cycle. (*5) Lateral peak is the abruptness of the shift from the upper lip to the lower lip on the vocal fold free edge (a parameter of vertical phase difference).

2.6. Digital Kymography

DKG analyzes mediolateral vocal fold movements at a selected longitudinal level. This method provides a vast quantitative capacity [9,12,16–18]. Single-line kymography (SLK) is generally advantageous for the assessment of temporal and mediolateral dynamics [17,18]. Intermittent longitudinal kinematics can also be assessed using multi-line kymography (MLK) [16].

The present study employed a five-line MLK to extract kymograms from the 10%, 30%, 50%, 70%, and 90% longitudinal levels of the glottal axis (a line between the vocal process and the anterior commissure). A kymogram of the 50% longitudinal level (the mid-membranous level) was used for the SLK parameters, and all five kymograms were used as MLK parameters.

Table 5. Parameters from digital kymography (DKG) 2: Multi-line kymography (MLK).

Parameter from MLK (N = 7)	Details
Continuous variables	
SI mean ^{MLK} (*1)	$= \dfrac{\sum_{i=0}^{5}\{lt.SI(i)+rt.SI(i)\}}{2\times5}$
SI difference ^{MLK}	$= \dfrac{\sum_{i=0}^{5}\{\mid lt.SI(i)+rt.SI(i)\mid\}}{5}$
OQ ^{MLK} (*2)	$= \dfrac{\sum_{i=0}^{5}OQ(i)}{5}$
N_G-opening long. PD mean ^{MLK} (*3)	(lt. opening long. PD + rt. opening long. PD)/GC
N_G-opening long. PD difference ^{MLK}	(lt. opening long. PD − rt. opening long. PD)/GC
N_G-closing long. PD mean ^{MLK} (*4)	(lt. closing long. PD + rt. closing long. PD)/GC
N_G-closing long. PD difference ^{MLK}	(lt. closing long. PD − rt. closing long. PD)/GC

SI = speed index; ^{MLK} = derived from multi-line kymography; lt. = left; rt. = right; OQ = open quotient; N_G- = normalized by glottal cycle; long. = longitudinal; PD = phase difference; GC = glottal cycle. All the MLK parameters are continuous variables. MLK continuous parameters are automatically calculated although some manual inputs are needed (e.g., to omit noise by secretion, to adjust length or area to measure). (*1) SI mean ^{MLK} signifies an average of all SIs (left and right SIs from 5 kymograms, 10 SI at a maximum). SI difference ^{MLK} signifies an average SI difference in 5 kymograms. (*2) OQ ^{MLK} signifies an average of OQs from 5 kymograms. (*3) Opening longitudinal phase difference is a temporal difference of glottal opening along the glottal axis. It is calculated by measuring the maximal temporal difference of glottal opening among 5 kymograms. (*4) Closing longitudinal phase difference is a temporal difference of glottal closure along the glottal axis. It is calculated by measuring the maximal temporal difference of glottal closure among 5 kymograms.

Table 6. Parameters from glottal area waveform (GAW).

Parameter from GAW (19)	Details
Continuous variables	
F_0 ^{GAW}	Hz
N_G-Opening Phase ^{GAW}	%, range: from 0 to 100
N_G-Closing Phase ^{GAW}	%, range: from 0 to 100
OQ ^{GAW}	Open phase/glottal cycle, range: from 0 to 1
SI ^{GAW}	(opening phase − closing phase)/(opening phase + closing phase), range: from −1 to 1
N_L-minimal GA ^{GAW}	Minimal GA/VFL^2
N_L-minimal GA flatness ^{GAW}	Minimal GA/(minimal GA circumference × VFL)
N_L-1/2O-GA ^{GAW}	GA at the midpoint of opening phase/VFL^2
N_L-1/2O-GA flatness ^{GAW}	GA at the midpoint of opening phase/(GA circumference at the midpoint of opening phase × VFL)
N_L-maximal GA ^{GAW}	Maximal GA/VFL^2
N_L-maximal GA flatness ^{GAW}	Maximal GA/(maximal GA circumference × VFL)
N_L-1/2C-GA ^{GAW}	GA at the midpoint of closing phase/VFL^2
N_L-1/2C-GA flatness ^{GAW}	GA at the midpoint of closing phase/(GA circumference at the midpoint of closing phase × VFL)
N_L-GA difference ^{GAW}	(maximal GA − minimal GA)/VFL^2
GA difference index ^{GAW}	(maximal GA − minimal GA)/maximal GA (range: from 0 to 1)
1/2O-Ratio ^{GAW}	GA at the midpoint of opening phase/maximal GA
1/2C-Ratio ^{GAW}	GA at the midpoint of closing phase/maximal GA
N_L-glottal outlet ^{GAW} (*1)	Glottal outlet/VFL^2
N_L-glottal outlet flatness ^{GAW}	Glottal outlet/(glottal outlet circumference × VFL)

F_0 = fundamental frequency; ^{GAW} = derived from glottal area waveform; N_G- = normalized by glottal cycle; OQ = open quotient; SI = speed index; N_L- = normalized by vocal fold length; GA = glottal area; VFL = vocal fold length; 1/2O- = at the midpoint of opening phase; 1/2C- = at the midpoint of closing phase. All the GAW parameters are continuous variables. GAW continuous parameters are automatically calculated although some manual inputs are needed (e.g., to omit noise by secretion, to adjust length or area to measure). (*1) Glottal outlet is a supraglottal area delineated by the false vocal folds, arytenoids, and epiglottis (a parameter of supraglottal hyperactivity).

From the SLK, 26 parameters were evaluated (Table 4): F0, open quotient (OQ), speed index (SI; mean and left-right difference), amplitude (mean and left-right difference), glottal width (maximum, integral value, left-right difference, and the ratio of the integral value), mucosal wave magnitude (distance of the visible mucosal wave traveling in the lateral direction; mean and difference), effective mucosal wave magnitude (mucosal wave

magnitude—amplitude; mean and left-right difference), mucosal wave persistence (mean and difference), effective mucosal wave persistence (mucosal wave persistence—opening phase; mean and left-right difference), lateral PD (magnitude and direction), axis shift (magnitude, ratio to amplitude and direction), and lateral peak (vertical PD; mean and left-right difference). These SLK-derived parameters were labeled with the suffix [SLK] (e.g., OQ^{SLK}).

From MLK, seven parameters were measured (Table 5): OQ, SI (mean and left-right difference), opening longitudinal PD (mean and left-right difference), and closing longitudinal PD (mean and left-right difference). The MLK-derived parameters were labeled with the suffix [MLK] (e.g., OQ^{MLK}).

2.7. Glottal Area Waveform

GAW provides information on the general dynamics of the glottal area by tracing the vocal fold edges using intensity-based threshold analysis to define the glottal area contour and display temporal changes in the glottal area [19–22]. Although GAW is a method for *t*-axis analysis alone, it has great potential for the quantitative analysis of vibratory dynamics by integrating *x*- and *y*-axis data of the glottal area into a scalar quantity [19].

This study evaluated 19 GAW parameters (Table 6): F0, opening phase, closing phase, OQ, SI, minimal glottal area (size and flatness), glottal area at the midpoint of the opening phase (size, flatness, and ratio to the maximal glottal area), maximal glottal area (size and flatness), glottal area at the midpoint of the closing phase (size, flatness, and ratio to the maximal glottal area), glottal area difference (size and ratio), and glottal area (supraglottal area; size and flatness). The GAW-derived parameters were labeled with the suffix [GAW] (e.g., OQ^{GAW}).

2.8. Videostroboscopy and Assessability

In patients with voice disorders, videostroboscopic examination was performed as a routine clinical assessment. To test the superiority of HSDI over VS, the assessability of vocal fold vibration was investigated. Assessability in this study was defined as whether relevant vibratory features could be extracted from the recorded sequence. Regarding VS, the video sequence of a sustained phonation at a comfortable pitch and loudness was selected.

2.9. Statistics

Chi-square tests were used to compare the assessability of vocal fold vibrations between the VS and HSDI.

For the univariate analyses between normal subjects and patients with laryngeal diseases, Kolmogorov–Smirnov tests were applied to assess the normal distribution of the parameters. Student's *t*-tests were performed for continuous parameters with a normal distribution. For the other parameters (e.g., categorical, non-normally distributed continuous parameters), Mann-Whitney U tests were performed. First, the normal subjects and all the patients were compared. Laryngeal diseases with more than 10 subjects were then separately compared to normal subjects.

Subsequently, binomial logistic regression analysis was applied to parameters with significant intergroup differences by univariate analysis. Calculations were performed by comparing the HSDI data (explanatory variable) of one laryngeal disease (objective variable) to those of normal subjects. In all analyses, the statistical significance was set at $p < 0.05$.

The coefficient of determination (R^2) is the degree of contribution of the parameters with significant differences to the characteristics of the subgroup. $R^2 \geq 0.7$, $0.7 > R^2 \geq 0.5$, $0.5 > R^2 \geq 0.3$, and $0.3 > R^2$ were considered high, moderate, low, and very low values, respectively [23].

The multicollinearity of the parameters was assessed using the variance inflation factor (VIF) according to the following formula (α = the correlation coefficient of two explanatory parameters), in which VIF < 5 indicated no multicollinearity.

$$VIF = 1/(1 - \alpha^2) \tag{1}$$

3. Results

3.1. All Laryngeal Disorders

In categorical parameters, the inter-rater reliability ranged from 83% to 100%, and the intra-rater reliability ranged from 90% to 100%.

The G score of the grade, roughness, breathiness, asthenia, and strain (GRBAS) scale of all patients was 1.78 ± 0.90. Visualization of vocal fold vibration was significantly more feasible by HSDI than by VS in the patient group (91% vs. 57%, $p < 0.001$).

Univariate analysis between normal subjects and all patients with voice disorders revealed 23 parameters with significant intergroup differences. Subsequent logistic regression analysis revealed four key parameters (Table 7): glottal gap size^{VPR}, atrophic change^{VPR}, N_G-effective mucosal wave persistence mean^{SLK}, and OQ^{MLK}. The R^2 value of these four parameters was very low (0.29) and there was no multicollinearity (the maximal VIF was 2.13 between the glottal gap size^{VPR} and OQ^{MLK}).

Table 7. Key parameters of all laryngeal disorders.

Explanatory Parameter (*1)	Normal (N = 46)	All Diseases (N = 304)	OR	*p* Value
Glottal gap size ^{VPR} [0, 3]	0.52	1.11	0.89	0.005 **
Atrophic change ^{VPR} [0, 3]	0.22	0.92	1.14	0.001 **
N_G-effective MWP mean ^{SLK} (%)	29.9 ± 18.7	13.6 ± 13.6	0.60	0.001 **
OQ ^{MLK} [0, 1]	0.51 ± 0.12	0.78 ± 0.21	2.28	<0.001 ***

OR = odds ratio; ^{VPR} = derived from visual perceptual rating; N_G- = normalized by glottal cycle; MWP = mucosal wave persistence; ^{SLK} = derived from single-line kymography; OQ = open quotient; ^{MLK} = derived from multi-line kymography; ** = $p < 0.01$; *** = $p < 0.001$. (*1) Parentheses and square brackets signify unit and range of parameter (e.g., [−3, 3] means −3 is lowest and +3 is highest). The range of parameters signified by % is from 0 to 100.

3.2. Vocal Fold Paralysis, Vocal Fold Polyp and Vocal Fold Scar

Univariate analysis between normal subjects and patients with vocal fold paralysis showed statistically significant differences in 26 parameters. Logistic regression analysis revealed eight key parameters (Table 8): periodicity^{LTG} (diprophonia), N_L-amplitude mean^{SLK}, N_G-mucosal wave persistence mean^{SLK}, OQ^{SLK}, OQ^{MLK}, SI^{SLK}, SI^{MLK}, and SI^{GAW}. The R^2 was high (0.76) and there was no multicollinearity (maximal VIF was 3.67 between OQ^{SLK} and OQ^{MLK}).

Comparisons between normal subjects and patients with vocal fold polyps showed statistically significant differences in 15 parameters. Logistic regression analysis revealed five key parameters (Table 8): periodicity^{LTG}, N_G-mucosal wave persistence mean^{SLK}, N_G-mucosal wave persistence difference^{SLK}, OQ^{MLK}, and N_G-opening longitudinal PD^{MLK}. The R^2 value was low (0.43) and there was no multicollinearity (the maximal VIF was 1.08 between the N_G-mucosal wave persistence mean^{SLK} and N_G-mucosal wave persistence difference^{SLK}).

Comparisons between normal subjects and patients with vocal fold scars showed statistically significant differences in 15 parameters. Subsequent logistic regression analysis revealed that only the non-vibrating area^{LTG} was significantly attributable to vocal fold scarring, with a moderate R^2 value of 0.68 (Table 8).

Table 8. Key parameters of vocal fold paralysis, polyp, and scarring.

Explanatory Parameter (*1)	Normal	Disease	OR	*p* Value
Vocal fold paralysis				
Periodicity ^{LTG} (%)	100	64	8.02	0.009 **
N_L-Amp. mean ^{SLK} (%)	7.6 ± 2.6	9.9 ± 4.1	113	0.036 *
N_G-MWP mean ^{SLK} (%)	54.3 ± 18.7	40.1 ± 11.2	0.09	<0.001 ***
OQ ^{SLK} [0, 1]	0.51 ± 0.13	0.84 ± 0.16	10.6	0.003 **
SI ^{SLK} [−1, 1]	-0.13 ± 0.19	-0.28 ± 0.15	7.1	0.002 **
OQ ^{MLK} [0, 1]	0.51 ± 0.12	0.89 ± 0.18	6.7	0.013 *
SI ^{MLK} [−1, 1]	-0.13 ± 0.19	-0.29 ± 0.19	0.23	0.012 *
SI ^{GAW} [−1, 1]	0.12 ± 0.17	-0.11 ± 0.14	0.11	0.012 *
Vocal fold polyp				
Periodicity ^{LTG} (%)	100	90	16.3	0.037 *
N_G-MWP mean ^{SLK} (%)	54.3 ± 18.7	43.1 ± 16.7	0.051	0.003 **
N_G-MWP diff. ^{SLK} (%)	13.5 ± 10.1	22.6 ± 20.3	77.1	0.001 **
N_G-OLPD ^{MLK} (%)	7.1 ± 24.9	-19.5 ± 26.3	0.12	0.001 **
OQ ^{MLK} [0, 1]	0.51 ± 0.12	0.63 ± 0.20	28.7	0.004 **
Vocal fold scar				
NVA ^{LTG} (%)	0	58	∞	<0.001 ***

OR = odds ratio; ^{LTG} = derived from laryngotopography; N_L- = normalized by vocal fold length; Amp. = amplitude; ^{SLK} = derived from single-line kymography; N_G- = normalized by glottal cycle; MWP = mucosal wave persistence; OQ = open quotient; SI = speed index; ^{MLK} = derived from multi-line kymography; ^{GAW} = derived from glottal area waveform; OLPD = opening longitudinal phase difference; NVA = non-vibrating area; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$. (*1) Parentheses and square brackets signify unit and range of parameter. The range of parameters signified by % is from 0 to 100.

3.3. Vocal Fold Atrophy and Sulcus Vocalis

Univariate analyses between normal subjects and patients with vocal fold atrophy showed statistically significant differences in 17 parameters. Subsequent logistic regression analysis revealed six key parameters (Table 9): N_G-lateral PD^{LTG}, N_{GL}-integral glottal width^{SLK}, OQ^{MLK}, SI^{MLK}, N_G-opening longitudinal PD^{MLK}, and N_{GL}-lateral peak index difference^{SLK}. The R^2 value was moderate (0.53) and there was no multicollinearity (the maximum VIF was 1.08 between the N_{GL}-integral glottal width^{SLK} and OQ^{MLK}).

Table 9. Key parameters of vocal fold atrophy and sulcus vocalis.

Explanatory Parameter (*1)	Normal	Disease	OR	*p* Value
Vocal fold atrophy				
N_G-lateral PD ^{LTG} (%)	3.9 ± 4.3	6.0 ± 4.1	8.5	0.015*
N_{GL}-IGW mean ^{SLK} (%)	4.2 ± 1.7	7.3 ± 3.3	240	<0.001***
N_{GL}-LPI difference ^{SLK} (%)	13.7 ± 13.4	24.4 ± 23.1	1.7	0.010*
OQ ^{MLK} [0, 1]	0.51 ± 0.12	0.67 ± 0.19	3.0	<0.001***
SI ^{MLK} [−1, 1]	-0.13 ± 0.19	-0.22 ± 0.15	0.61	0.023*
N_G-OLPD ^{MLK} (%)	7.2 ± 25.1	-7.9 ± 21.0	0.62	0.003**
Sulcus vocalis				
N_G-lateral PD ^{LTG} (%)	3.9 ± 4.3	7.3 ± 6.8	3.4	0.024*
N_{GL}-IGW mean ^{SLK} (%)	4.2 ± 1.7	10.3 ± 4.9	30.1	<0.001***
OQ ^{MLK} [0, 1]	0.51 ± 0.13	0.82 ± 0.15	4.4	<0.001***
SI ^{MLK} [−1, 1]	-0.13 ± 0.19	-0.25 ± 0.15	0.51	0.014*

OR = odds ratio; N_G- = normalized by glottal cycle; PD = phase difference; ^{LTG} = derived from laryngotopography; N_{GL}- = normalized by glottal cycle and vocal fold length; IGW = integral glottal width; ^{SLK} = derived from single-line kymography; LPI = lateral peak index; OQ = open quotient; ^{MLK} = derived from multi-line kymography; SI = speed index; OLPD = opening longitudinal phase difference; * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$. (*1) Parentheses and square brackets signify unit and range of parameter. The range of parameters signified by % is from 0 to 100.

Comparisons between normal subjects and patients with sulcus vocalis showed statistically significant differences in 18 parameters. Logistic regression analysis revealed four key parameters (Table 9): N_G-lateral PD^{LTG}, N_{GL}-integral glottal width^{SLK}, OQ^{MLK}, and

SI^{MLK}. The R^2 value was high (0.74) and there was no multicollinearity (the maximal VIF was 1.45 between the N_{GL}-integral glottal width^{SLK} and OQ^{MLK}).

3.4. Laryngeal Leukoplakia and Laryngeal Cancer

Comparison between normal subjects and patients with laryngeal leukoplakia showed statistically significant differences in eight parameters. Logistic regression analysis revealed that N_G-lateral PD^{LTG} and N_{GL}-integral glottal width^{SLK} were key parameters (Table 10). The R^2 value was low (0.42) and there was no multicollinearity (VIF = 1.12).

Table 10. Key parameters of laryngeal leukoplakia and cancer.

Explanatory Parameter	Normal	Disease	OR	*p* Value
Laryngeal leukoplakia				
N_G-lateral PD ^{LTG} (%)	3.9 ± 4.3	7.9 ± 8.0	2.9×10^6	0.003 **
N_{GL}-IGW mean ^{SLK} (%)	4.2 ± 1.7	8.5 ± 3.1	3.0×10^{20}	<0.001 ***
Laryngeal cancer				
NVA ^{LTG} (%)	0	50	∞	<0.001 ***

OR = odds ratio; N_G- = normalized by glottal cycle; PD = phase difference; ^{LTG} = derived from laryngotopography; N_{GL}- = normalized by glottal cycle and vocal fold length; IGW = integral glottal width; ^{SLK} = derived from single-line kymography; NVA = non-vibrating area; ** = $p < 0.01$; *** = $p < 0.001$. The range of parameters signified by % is from 0 to 100.

Univariate analyses between normal subjects and patients with laryngeal cancer showed statistically significant differences in the four parameters. Logistic regression analysis revealed that only the non-vibrating area was significantly attributable to laryngeal cancer, with a moderate R^2 value (0.52; Table 10).

3.5. Summary of Key Parameters

The comparison of normal subjects and all patients or patients from each laryngeal disorder revealed a total of 17 key parameters with duplication omitted: by method, eight were from SLK, three were from MLK, three were from LTG, two were from VPR, and one was from GAW. By parameter type, 11, 4, and 2 were the numbers of time, size, and size-and-time parameters, respectively.

4. Discussion

4.1. VS versus HSDI

The results of the present study demonstrated the superiority of HSDI to VS. First, the multifaceted assessment was feasible owing to the combination of four analysis methods in HSDI, while VS is basically assessed by VPR [3]. Second, visualization of vocal fold vibration was more frequently possible by HSDI (91%) than by VS (57%) in the patient group. This difference was presumably due to moderate-to-severe dysphonia in the patient group, as demonstrated by the average G score of 1.78: stroboscopic light tracking and frame rate cannot follow highly aperiodic voices, leading to tracking errors. Patel et al. reported rates of successful vibratory assessment in subjects with G1, G2, and G3 voices of 100%, 36%, and 0% by VS, compared to 100% by HSDI [2].

Furthermore, as 76% of the key parameters (13/17) were related to the time domain, HSDI, which has a high temporal resolution (4500 fps in the present study), is theoretically advantageous over VS (30–60 fps).

4.2. Analysis Method

While VPR is the gold standard for VS, there is no standardized analysis method for HSDI [1,5]. Thus, the authors chose the analysis methods based on our familiarity with the technique, and on the capability of the method to analyze three dimensions of HSDI data: lateral (*x*-axis), longitudinal (*y*-axis), and temporal (*t*-axis) data (Figure 1).

VPR is effective in the analysis of *x*-axis and *y*-axis data but not of *t*-axis data since this method is based on frame-by-frame or slow-motion playback of HSDI data, and as complex, moderate-to-long-term temporal dynamics such as inter-cycle irregularity are difficult to assess [13]. LTG is effective for objective, quantitative, *x*-axis, and *y*-axis evaluations [14,15]. DKG is effective in *x*- and *t*-axis data analysis, as well as in *y*-axis data, to a certain extent, when MLK is used [16,17]. Although GAW is a method for t-axis analysis alone, it has great potential for the quantitative analysis of vibratory dynamics by integrating *x*- and *y*-axis data of the glottal area into a scalar quantity [19].

Although each method is not sufficiently omnipotent, the combination of these techniques allows for sufficient analysis of all three dimensions of the HSDI data. PVG may be added to the current analysis program because it is considered to be effective in the y- and t-axis data and thus complements the HSDI analysis by DKG (*x*- and *t*-axes) and LTG (*x*- and *y*-axes) [6,7].

DKG accounted for the largest proportion of key parameters (67%, 11/17) in this study, with 33% of the DKG parameter list (11/33; Tables 4 and 5). Although we identified only three LTG-derived key parameters, this corresponded to 30% of the LTG parameter set (3/10; Table 3); thus, DKG and LTG were considered equivalent in this study.

VPR, with two key parameters (9% of the 21-parameter set), played a less important role. Thus, VPR, based on the examiners' subjective judgment, may not be effective enough to judge complex dynamics. Similarly, GAW accounted for one key parameter (5% of the 19-GAW-parameter list) and contributed less to this study. Given that the GAW parameter lists differed between the present study and Bohr's works, implementations of parameters not attempted in the present work but included in Bohr's studies (e.g., maximum area declination rate, shimmer, jitter) may lead to different results [6,7].

4.3. Parameters

In this work, some measures (e.g., OQ, SI, and F_0) are computed using different methods. There are two reasons for this. First, methods to compute a certain measure can differ among different studies, and thus assessing measures by different methods can facilitate comparisons with the literature.

Second, the characteristics of the measures change depending on the computed methods, and thus it is worthwhile to test the same measures derived from different calculations [24]. For instance, Yokonishi et al. compared various OQs derived from EGG, GAW, SLK, MLK, and kymography averaged along the glottal axis, and reported that OQs from kymography averaged along the glottal axis and OQ^{MLK} revealed the strongest correlation with acoustic property recorded simultaneously with the HSDI, and these OQs best described changes in phonation types [24]. Similarly, in the current study, OQ^{MLK} was found to be a key parameter in various laryngeal diseases, unlike OQ^{SLK} and OQ^{GAW}, presumably because OQ^{MLK} is more sensitive to abnormalities in vocal fold vibrations and can better represent the vibratory characteristics of laryngeal disorders than the other two.

Caution is needed as analyzing similar parameters in multivariate analysis can lead to the risk of multicollinearity. However, no multicollinearity among key parameters was found in the current study, and thus that potential problem was not detected in this study.

4.4. All Laryngeal Disorders

The R^2 value was lower for laryngeal disorders as a whole (0.29) than for each laryngeal disease (0.42–0.76). First, the implementation of parameters that should better reflect general vibratory abnormalities (e.g., cepstral peak prominence, entropy) may improve the analysis [25,26]. Second, this result suggests that the key parameters differ for each laryngeal disorder, and thus should be assessed disease-specifically rather than holistically. However, the R^2 values varied widely. While the current approach was construed as effective in diseases with moderate-to-high R^2 values, exploration of more specific parameters is needed in laryngeal disorders with low-to-very low R^2 values.

4.5. Vocal Fold Paralysis

Considering the high R^2 value for vocal fold paralysis, eight key parameters were selected to sufficiently explain the general vibratory characteristics of this clinical entity. Periodic irregularity may result from the decreased coupling of the left and right vocal folds due to geometric separation of the vocal folds by the paralysis, or due to the asymmetry of material properties between the left and right vocal folds [11,27,28]. High OQs should reflect glottal insufficiency, while low SIs (opening phase < closing phase) should reflect reduced vocal fold tension or mass by denervation or muscular atrophy [11,27,28]. The increased amplitude may be due to reduced mass and tension in paralyzed vocal folds or a compensatory increase in subglottal pressure, while reduced mucosal waves can result from reduced energy transfer to the mucosa due to glottal insufficiency [11,27,28]. Caution is needed in that there is literature inconsistency regarding the amplitude and mucosal wave of vocal fold paralysis, which is presumably related to differences in phonatory condition (e.g., subglottal pressure, pitch), material properties (mass, tension), or anatomical condition (paralyzed position, vertical level) [29–31].

4.6. Vocal Fold Atrophy and Sulcus Vocalis

There were similarities in the key parameters between vocal fold atrophy and sulcus vocalis, although the degree of abnormalities was generally more severe in sulcus vocalis. The greater severity of vibratory disturbance in the sulcus vocalis was consistent with reports in the literature [32–35]. However, the vibratory similarity may stem from a certain extent of pathogenetic overlap between these diseases: sulcus vocalis can be divided into three subtypes (type 1, physiological sulcus; type 2, sulcus vergeture; and type 3, open cyst), with type 1 sulcus considered an age-related phenomenon such as vocal fold atrophy [32–35].

Lateral PDs should result from left-right asymmetry in the material properties of the vocal fold lamina propria due to geriatric changes in the intralaryngeal muscles or sulci [32–35]. Greater N_{GL}-integral glottal width[SLK] and larger OQ[MLK] should reflect impaired glottal closure, while lower SIs should reflect reduced mass or tension of the vocal folds due to geriatric changes in the laryngeal muscles and lamina propria or sulcal changes in the lamina propria [32–35]. Longitudinal and vertical PDs observed in vocal fold atrophy can stem from asymmetry in the mass, tension, mucoelasticity, and glottal closure of the vocal folds due to geriatric changes of the glottis in the longitudinal or vertical direction [8,12,18].

4.7. Laryngeal Leukoplakia and Cancer

Classically reported vibratory features were normal in laryngeal leukoplakia and the presence of non-vibrating areas in laryngeal cancer [36–38]. The present study generally coincided with these classical findings: while N_G-lateral PD[LTG] and N_{GL}-integral glottal width[SLK] were the key parameters in laryngeal leukoplakia, non-vibrating area was the sole key parameter in laryngeal cancer.

4.8. Vocal Fold Polyp

The vibratory characteristics of vocal fold polyps were frequent periodic irregularity, reduced N_G-mucosal wave persistence mean[SLK], greater N_G-mucosal wave persistence difference[SLK], larger OQ[MLK], and larger anterior-to-posterior N_G-opening longitudinal PD[MLK]. This can be explained by the left-right asymmetry of the material property between the sides with or without polyps, or by the glottal insufficiency resulting from the entrapment of a polyp between the vocal fold free edges [39,40]. However, the R^2 values of these key parameters were low (0.43). Thus, there may be better, yet unexplored, parameters for this clinical entity.

4.9. Vocal Fold Scar

In vocal fold scars, the non-vibrating area^{LTG} was the sole key parameter, and R^2 was moderate. Thus, the detection of non-vibrating areas is important for this clinical entity. Although non-vibrating areas were also primarily important in laryngeal cancer, consideration of other factors such as the presence of an evident organic tumor in laryngeal cancer and the absence of vocal fold scarring may lead to discrimination of the two clinical entities [36,41,42].

4.10. Limitations

To the best of our knowledge, this is the first study to clarify the key parameters of vocal fold vibration in a disease-specific manner by multivariate analyses, and the obtained results were consistent with those previously reported and considered to be reasonable. Further, the current study analyzed the largest subject population and parameter set, as well as the largest number of analysis methods on this topic. All the methods provided key parameters, and thus the selection of the method was considered to be sufficient. Thus, the strategy of utilizing multiple methods in the current work was considered to be effective, and the implementation of methods that are effective in the analysis of the lateral direction, such as DKG and LTG, is recommended for the assessment of abnormal vocal fold vibration. Thus, the current study can serve as a reference for normal and abnormal vocal fold vibrations and offer a guide for which parameters to select in relevant future HSDI studies.

Nevertheless, there are limitations: First, there was a variation in subject size among disorders and a lack of major laryngeal disorders such as vocal fold nodules, vocal fold cysts, and functional disorders. An unbalanced sample size among voice disorders is another limitation: a wide range of sample sizes may affect the statistical results. Increasing the subject number, adding other major laryngeal disorders, and balancing the sample sizes will be required in future studies.

Prematurity and incompletion of the analysis techniques or parameters could be another limitation. Although the present study utilized the largest analysis methods on this topic, PVG, which is effective in the analysis of y-axis and t-axis data, may reinforce the analysis potential. Although the parameter set in the current work is the largest as far as we know, the list should be updated by adding new parameters or parameters that are already present but not tried in the present work. Additionally, refining analysis techniques will be needed to promote automation and decrease the necessity of user input.

Finally, the results of the current study should be applied to sustained phonation with a comfortable pitch and loudness because dysphonia can be better detected during nonstationary phonation or during more vocally demanding tasks than during sustained phonation. The evaluation of other phonation tasks (e.g., vocal onset, high-pitch/low-pitch phonation) is needed in the future.

5. Conclusions

In the future, HSDI will be the gold standard for vocal fold vibratory assessment because of its potential for multifaceted, objective, and quantitative analyses. However, HSDI is not yet widely used in clinical practice due to unsolved disadvantages: the relative scarcity of studies to prove the merits of HSDI and its superiority to VS, high cost, time consumed during data processing and analysis, and the lack of standardized techniques and parameters. Many researchers involved in HSDI try to solve these disadvantages, and the current study is one of these trials. The current work demonstrates the superiority of HSDI to VS by showing better assessment of vocal fold vibrations and the capacity of multifaceted vibratory analysis using various analysis methods. Furthermore, an example of both subjective and objective qualitative and quantitative analysis methods using a large parameter set was provided, which successfully differentiated normal and abnormal vocal fold vibrations, and clarified disease-specific key parameters. DKG and LTG were particularly effective in the evaluation of abnormal vocal fold vibrations, and thus the im-

plementation of methods effective in the analysis of lateral direction data is recommended for the assessment of abnormal vocal fold vibration. Although there are limitations to be solved in the future, this study is considered to be a new step toward evidence-based medicine of voice disorders.

References

1. Cohen, S.M.; Pitman, M.J.; Noordzij, J.P.; Courey, M. Evaluation of dysphonic patients by general otolaryngologists. *J. Voice* **2012**, *26*, 772–778. [CrossRef] [PubMed]
2. Patel, R.R.; Dailey, S.; Bless, D.M. Comparison of high-speed digital imaging with stroboscopy for laryngeal imaging of glottal disorders. *Ann. Otol. Rhinol. Laryngol.* **2008**, *117*, 413–424. [CrossRef] [PubMed]
3. Mehta, D.D.; Hillman, R.F. Current role of stroboscopy in laryngeal imaging. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2012**, *20*, 429–436. [CrossRef]
4. Hertegard, S. What have we learned about laryngeal physiology from high-speed digital videoendoscopy? *Curr. Opin. Otolaryngol. Head Neck Surg.* **2005**, *13*, 152–156. [CrossRef] [PubMed]
5. Deliyski, D.D.; Petrushev, P.P.; Bonilha, H.S.; Gerlach, T.T.; Martin-Harris, B.; Hillman, R.E. Clinical implementation of laryngeal high-speed videoendoscopy: Challenges and evolution. *Folia Phoniatr. Logop.* **2008**, *60*, 33–44. [CrossRef]
6. Bohr, C.; Kräck, A.; Eysholdt, U.; Ziethe, A.; Döllinger, M. Quantitative analysis of organic vocal fold pathologies in females by high-speed endoscopy. *Laryngoscope* **2012**, *123*, 1686–1693. [CrossRef]
7. Bohr, C.; Kräck, A.; Dubrovskiy, D.; Eysholdt, U.; Svec, J.; Psychogios, G.; Ziethe, A.; Döllinger, M. Spatiotemporal analysis of high-speed videolaryngoscopic imaging of organic pathologies in males. *J. Speech Lang. Heart Res.* **2014**, *57*, 1148–1161. [CrossRef]
8. Yamauchi, A.; Yokonishi, H.; Imagawa, H.; Sakakibara, K.I.; Nito, T.; Tayama, N.; Yamasoba, T. Vocal fold vibration in vocal fold atrophy: Quantitative analysis with high-speed digital imaging. *J. Voice* **2015**, *29*, 755–762. [CrossRef]
9. Yamauchi, A.; Yokonishi, H.; Imagawa, H.; Sakakibara, K.I.; Nito, T.; Tayama, N.; Yamasoba, T. Quantification of vocal fold vibration in various laryngeal disorders using high-speed digital imaging. *J. Voice* **2016**, *30*, 205–214. [CrossRef]
10. Yamauchi, A.; Yokonishi, H.; Imagawa, H.; Sakakibara, K.I.; Nito, T.; Tayama, N.; Yamasoba, T. Visualization and estimation of vibratory disturbance in vocal fold scar using high-speed digital imaging. *J. Voice* **2016**, *30*, 493–500. [CrossRef]
11. Yamauchi, A.; Yokonishi, H.; Imagawa, H.; Sakakibara, K.I.; Nito, T.; Tayama, N.; Yamasoba, T. Quantitative analysis of vocal fold vibration in vocal fold paralysis with the use of high-speed digital imaging. *J. Voice* **2016**, *30*, e13–e22. [CrossRef]
12. Yamauchi, A.; Yokonishi, H.; Imagawa, H.; Sakakibara, K.-I.; Nito, T.; Tayama, N.; Yamasoba, T. Characterization of vocal fold vibration in sulcus vocalis using high-speed digital imaging. *J. Speech Lang. Heart Res.* **2017**, *60*, 24–37. [CrossRef] [PubMed]
13. Yamauchi, A.; Imagawa, H.; Yokonishi, H.; Nito, T.; Yamasoba, T.; Goto, T.; Takano, S.; Sakakibara, K.-I.; Tayama, N. Evaluation of vocal fold vibration with an assessment form for high-speed digital imaging: Comparative study between healthy young and elderly subjects. *J. Voice* **2012**, *26*, 742–750. [CrossRef] [PubMed]
14. Yamauchi, A.; Imagawa, H.; Sakakibara, K.-I.; Yokonishi, H.; Nito, T.; Yamasoba, T.; Tayama, N. Phase difference of vocally healthy subjects in high-speed digital imaging analyzed with laryngotopography. *J. Voice* **2013**, *27*, 39–45. [CrossRef] [PubMed]
15. Granqvist, S.; Lindestad, P.A. A method of applying Fourier analysis to high-speed laryngoscopy. *J. Acoust. Soc. Am.* **2001**, *110*, 3193–3197. [CrossRef] [PubMed]
16. Yamauchi, A.; Imagawa, H.; Sakakibara, K.-I.; Yokonishi, H.; Nito, T.; Yamasoba, T.; Tayama, N. Characteristics of vocal fold vibrations in vocally healthy subjects: Analysis with multi-line kymography. *J. Speech Lang. Heart Res.* **2014**, *57*, S648–S657. [CrossRef]

17. Yamauchi, A.; Imagawa, H.; Sakakibara, K.-I.; Yokonishi, H.; Nito, T.; Yamasoba, T.; Tayama, N. Quantitative analysis of digital videokymography: A preliminary study on age- and gender-related difference of vocal fold vibration in normal speakers. *J. Voice* **2015**, *29*, 109–119. [CrossRef] [PubMed]

18. Svec, J.G.; Sram, F.; Schutte, H.K. Videokymography in voice disorders: What to look for? *Ann. Otol. Rhinol. Laryngol.* **2007**, *116*, 172–180. [CrossRef] [PubMed]

19. Yamauchi, A.; Imagawa, H.; Sakakibara, K.-I.; Yokonishi, H.; Nito, T.; Yamasoba, T.; Tayama, N. Age- and gender-related difference of vocal fold vibration and glottal configuration in normal speakers: Analysis with glottal area waveform. *J. Voice* **2014**, *28*, 525–531. [CrossRef]

20. Bloch, I.; Behrman, A. Quantitative analysis of videostroboscopic images in presbylarynges. *Laryngoscope* **2001**, *111*, 2022–2027. [CrossRef] [PubMed]

21. Inwald, E.C.; Döllinger, M.; Schuster, M.; Eysholdt, U.; Bohr, C. Multiparametric analysis of vocal fold vibrations in healthy and disordered voices in high-speed imaging. *J. Voice* **2011**, *25*, 576–590. [CrossRef] [PubMed]

22. Mehta, D.D.; Delyski, D.D.; Quatieri, T.F.; Hillman, R.F. Automated measurement of vocal fold vibratory asymmetry from high-speed videoendoscopy recordings. *J. Speech Heart Res.* **2011**, *54*, 47–54. [CrossRef]

23. Moore, D.S.; Notz, W.I.; Flinger, M.A. *The Basic Practice of Statistics*, 6th ed.; W. H. Freeman and Company: New York, NY, USA, 2013; p. 138.

24. Yokonishi, H.; Imagawa, I.; Sakakibara, K.-I.; Yamauchi, A.; Nito, N.; Yamasoba, T.; Tayama, N. Relationship of Various Open Quotients With Acoustic Property, Phonation Types, Fundamental Frequency, and Intensity. *J. Voice* **2016**, *30*, 145–157. [CrossRef]

25. Mehta, D.D.; Zeitels, S.M.; Burns, J.A.; Friedman, A.D.; Deliyski, D.D.; Hillman, R.E. High-speed videoendoscopic analysis of relationships between cepstral-based acoustic measures and voice production mechanisms in patients undergoing phonomicrosurgery. *Ann. Otol. Rhinol. Laryngol.* **2012**, *121*, 341–347. [CrossRef]

26. Wang, J.S.; Olszewski, E.; Devine, E.E.; Hoffman, M.R.; Zhang, Y.; Shao, J.; Jiang, J.J. Extension and application of high-speed digital imaging analysis via spatiotemporal correlation and eigenmode analysis of vocal fold vibration before and after polyp excision. *Ann. Otol. Rhinol. Laryngol.* **2016**, *125*, 660–666. [CrossRef] [PubMed]

27. Rubin, A.D.; Sataloff, R.T. Vocal fold paresis and paralysis. *Otolaryngol. Clin. N. Am.* **2007**, *40*, 1109–1131. [CrossRef]

28. Misono, S.; Merati, A.L. Evidence-based practice: Evaluation and management of unilateral vocal fold paralysis. *Otolaryngol. Clin. N. Am.* **2012**, *45*, 1083–1108. [CrossRef]

29. Schutte, H.K.; Svec, J.G.; Sram, F. First results of clinical application of videokymography. *Laryngoscope* **1998**, *108*, 1206–1210. [CrossRef]

30. Voigt, D.; Döllinger, M.; Eysholdt, U.; Yang, A.; Gürlek, E.; Lohscheller, J. Objective detection and quantification of mucosal wave propagation. *J. Acoust. Soc. Am.* **2010**, *128*, 347–355. [CrossRef]

31. Qiu, Q.; Schutte, H.K. A new generation videokymography for routine clinical vocal fold examination. *Laryngoscope* **2006**, *116*, 1824–1828. [CrossRef] [PubMed]

32. Kendall, K. Presbyphonia: A review. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2007**, *15*, 137–140. [CrossRef]

33. Yamauchi, A.; Imagawa, H.; Sakakibara, K.-I.; Yokonishi, H.; Nito, T.; Tayama, N.; Yamasoba, T. Vocal fold atrophy in a Japanese tertiary medical institute: Status quo of the most aged country. *J. Voice* **2014**, *28*, 231–236. [CrossRef]

34. Ford, C.N.; Inagi, K.; Khidr, A.; Bless, D.M.; Gilchrist, K.W. Sulcus vocalis: A rational analytical approach to diagnosis and management. *Ann. Otol. Rhinol. Laryngol.* **1996**, *105*, 189–200. [CrossRef]

35. Giovanni, A.; Chanteret, C.; Lagier, A. Sulcus vocalis: A review. *Eur. Arch. Otorhinolaryngol.* **2007**, *264*, 649–652. [CrossRef]

36. Colden, D.; Jarboe, J.; Zeitels, S.M.; Bunting, G.; Hillman, R.E.; Spanou, K. Stroboscopic assessment of vocal fold keratosis and glottic cancer. *Ann. Otol. Rhinol. Laryngol.* **2001**, *110*, 293–298. [CrossRef]

37. Mehta, D.D.; Deliyski, D.D.; Zeitels, S.M.; Quatieri, T.F.; Hillman, R.E. Voice production mechanisms following phonosurgical treatment of early glottic cancer. *Ann. Otol. Rhinol. Laryngol.* **2010**, *119*, 1–9. [CrossRef]

38. Djukic, V.; Milovanovic, J.; Jotic, A.D.; Vukasinovic, M. Stroboscopy in detection of laryngeal dysplasia effectiveness and limitations. *J. Voice* **2014**, *28*, 13–21. [CrossRef] [PubMed]

39. Chodara, A.M.; Krausert, C.R.; Jiang, J.J. Kymographic characterization of vibration in human vocal folds with nodules and polyps. *Laryngoscope* **2012**, *122*, 58–65. [CrossRef] [PubMed]

40. Kunduk, M.; Döllinger, M.; McWhorter, A.J.; Svec, J.G.; Lohscheller, J. Vocal fold vibratory behavior changes following surgical treatment of polyps investigated with high-speed videoendoscopy and phonovibrography. *Ann. Otol. Rhinol. Laryngol.* **2012**, *121*, 355–363. [CrossRef] [PubMed]

41. Bless, D.M.; Welham, N.V. Characterization of vocal fold scar formation, prophylaxis and treatment using animal models. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2010**, *18*, 481–486. [CrossRef] [PubMed]

42. Friedrich, G.; Dikkers, F.G.; Arens, C.; Remacle, M.; Hess, M.; Giovanni, A.; Duflo, S.; Hantzakos, A.; Bachy, V.; Gugatschka, M. Vocal fold scars: Current concepts and future directions. Consensus report of the phonosurgery committee of the European laryngological society. *Eur. Arch. Otorhinolaryngol.* **2013**, *270*, 2491–2507. [CrossRef] [PubMed]