Phonation Quotient Using Three Aerodynamic Instruments in the Disordered Voice

[a]Ashwini Joshi, [b]Christopher R. Watts, and [c]Jeremy Hathway, ac*Houston, and* b*Fort Worth, Texas*

Summary: Objective. The purpose of this study was to examine measures of phonation quotient (PQ) in two groups of persons with voice disorders using three different aerodynamic instruments representing low-tech and high-tech options.

Study Design. Prospective, repeated measures design.

Methods. Two groups of patients with a diagnosis of vocal fold paralysis/paresis ($n = 9$) or benign vocal fold lesions ($n = 8$) were assessed for maximum phonation time and vital capacity obtained from three aerodynamic instruments: a hand-held analog windmill-type spirometer; a hand-held digital spirometer; and the Phonatory Aerodynamic System (PAS), Model 6600. PQ was calculated using vital capacity from each instrument along with maximum phonation time. Univariate ANOVAs were performed to test for the main effects of disorder and instrument on derived PQ. Paired samples *t* tests were performed *post hoc* to investigate any significant main effects. Pearson product-moment correlation was performed to assess measurement reliability (parallel forms) between the instruments.

Results. Statistically significant differences were found for measures of PQ as a function of disorder but not instrument type. There was not a significant interaction effect between disorder and instrument type. Strong positive correlations were present between all three instruments for measures of PQ.

Conclusions. PQ was sensitive to differences in airflow as a function of disorder etiology (paralysis/paresis vs benign lesions). This clinical measure of laryngeal function can be used as a low-cost substitute in the absence of a pneumotachograph. These results are consistent with previous literature reporting data from adult male and female speakers with normal voice quality, and support the use of low-tech options for measurement of basic aerodynamic variables associated with voice production.

Key Words: Spirometer—Aerodynamics—Phonation quotient—Vocal fold paralysis—Benign lesions.

INTRODUCTION

Production of normal voice quality is dependent on the health of the individual subsystems of respiration, phonation, and resonance as well as the interaction of these systems.[1] It is, therefore, necessary to be able to assess these systems in isolation along with their resultant interactions. Aerodynamic assessment evaluates the efficiency of the valving mechanism of the larynx to the air coming through the trachea, thus evaluating laryngeal and respiratory function. It provides information about the ability of the laryngeal complex to control the airflow and maintain vocal fold vibration for normal voice production. As the power source of the voice, it is important to understand the functioning of the respiratory system for voice production and that its assessment is an important component of the clinical voice evaluation. It provides the clinician with the tools for a more accurate diagnosis, effective biofeedback, and as a meaningful measurement of change with treatment.[2]

Clinical evaluation of aerodynamics involves measuring volume, rate of airflow, subglottal pressure, and voicing efficiency.[3] Volume is measured by gathering data on vital capacity (VC), phonation volume, and forced expiratory volume. Rate of airflow is estimated using measures of peak or mean expiratory airflow. Subglottal pressure and phonation threshold pressure provide an estimate of tracheal air pressure, while estimates of vocal efficiency are calculated using maximum phonation time (MPT), aerodynamic efficiency, *s*/*z* ratio, or phonation quotient (PQ).[4-7] These assessments can be performed with a range of instrumentation using a basic stopwatch for MPT and *s*/*z* ratio to a complex pneumotachograph for measures of volume, flow, and pressure. While MPT and *s*/*z* ratio are used frequently by clinicians due to their ease of measurement and limited need for instrumentation, other measures of flow, volume, and pressure are often not part of the diagnostic protocol. Beginning clinicians, clinicians not specialized in voice disorders or those with limited resources may miss out on obtaining these aerodynamic measures, a critical piece of diagnostics, if they do not have precision high-tech equipment. However, clinicians have the option of using low-cost equipment such as a timer and a spirometer to derive aerodynamic measures. These do not allow for pressure measurements, but do provide physiological data on the status of the systems. Clinicians can obtain measurements for MPT with a digital timer, and VC with a spirometer to derive PQ defined as the ratio of largest VC to longest MPT,[4] to provide an indirect estimate of glottal airflow.

The reliability of PQ as a measure was first reported by Hirano et al[4] who found a strong correlation between

Accepted for publication August 2, 2018.

From the [a]University of Houston, Houston, Texas; [b]Texas Christian University, Fort Worth, Texas; and the [c]Texas Voice Center, Houston Methodist, Houston, Texas.

Address correspondence and reprint requests to Ashwini Joshi, Department of Communication Sciences and Disorders, University of Houston, 4455 Cullen Blvd, Houston, TX 77204. E-mail: ajoshi4@uh.edu

Journal of Voice, Vol. 34, No. 1, pp. 20—24

0892-1997

https://doi.org/10.1016/j.jvoice.2018.08.002

pneumotachograph-based mean flow rate measures and PQ measures obtained with a low-tech instrument. In our prior work, we demonstrated a strong correlation between the pneumotachograph and two spirometers in healthy adults.[8-10] PQ has been used as a diagnostic tool in the disordered voice population for individuals with benign and malignant lesions, functional disorders, vocal fold inflammation, and neurologic voice disorders.[4,5,11-17] Individuals with a disordered voice secondary to glottal incompetence have reduced MPT and hence have higher PQ values than those with normal voice quality.[4] Airflow rates for healthy individuals fall across a wide range[18] and hence the higher PQ values in the disordered voice are not always indicative of poor aerodynamic function. It does, however, give an insight into individual functioning and can be used to monitor change over time.

One of the difficulties with using PQ is the limited research on the reliability and validity of contemporary spirometers in comparison to high-tech instruments typically used in clinical practice. The precision of low-tech handheld spirometers likely varies depending on the type and quality of manufacturing. Choosing a cost-effective but precise spirometer can be difficult in the absence of research studies demonstrating the reliability of a specific type of spirometer and comparing it to a standard clinical instrument such as a pneumotachograph. Parallel forms or alternative forms reliability needs to be assessed for spirometers where the same variable is measured using different forms/versions of an instrument[19] to determine if the two instruments produce equivalent results under identical measurement conditions. Our previous publications have demonstrated strong parallel forms reliability between an analog (Baseline windmill), digital (Contec SP10), and a pneumotachograph (PENTAX Medical Phonatory Aerodynamic System [PAS]) across ages and sex in healthy adults.[8-10] The purpose of this study was to investigate the parallel forms measurement reliability and validity between the same three commercially available instrument systems for individuals with a disordered voice quality and glottal incompetence to assess if the results from the normative study can be generalized to the patient population. The study protocol was conducted in two groups of patients: those with glottal incompetence from vocal fold immobility (unilateral vocal fold paralysis/paresis), and those with incompetence from a benign vocal fold lesion to examine if PQ, a measure of airflow, is viable for those known to have a significant disruption to glottal airflow.

METHODS

Participants

Adult participants from a clinical voice practice were recruited in each of the two disordered voice groups—benign lesion ($n = 8$) or unilateral vocal fold paralysis ($n = 9$). The benign lesion group had seven females and one male, ranging in age from 19 to 61 years. The unilateral vocal fold paralysis group had six females and three males, ranging in age from 23 to 79 years. Effect sizes calculated based on previously published data on PQ were predicted to be large (eg, >0.85). The total sample size required to find a significant effect at alpha = 0.05 was $n = 6$ participants in each group. All participants demonstrated abnormal voice quality on perceptual evaluation by an experienced speech language pathologist and glottal incompetence on laryngeal endoscopy. The participants did not have a history of cognitive impairment or hearing loss. Participant information is provided in Table 1. The study was approved by the Institutional Review Board at the University of Houston.

Instruments

Two spirometers and a pneumotachograph-based system were used to collect measures of VC, from which the variable of interest (PQ) was derived. A windmill-type (Baseline Measurement Instruments, Fabrication Enterprises, Inc., White Plains, NY) and a digital spirometer (SP10, Contec Medical, China) were the hand-held spirometers used, while the PAS Model 6600 (KayPENTAX Corp, Lincoln Park, NJ)[20] was used as the pneumotachograph-based system. The analog windmill-type spirometer functions by using airflow through the mouthpiece to move the dial around a measurement window on the spirometer. The digital SP10 spirometer also includes a mouthpiece through which air is blown against the resistance of internal metal blades. The PAS includes custom software to digitize and process the aerodynamic signal. The PAS was calibrated prior to each measurement session according to the directions in the instruction manual.[20]

TABLE 1.
Demographic Information for Participants in the Two Disorder Groups

Age	Sex	Diagnosis
Paralysis/paresis		
65	Female	Left paralysis
64	Female	Left paralysis
70	Female	Right paralysis
61	Female	Right paralysis
59	Female	Left paralysis
33	Female	Right paralysis
72	Male	Right paresis
79	Male	Left paralysis
23	Male	Severe left paresis
Benign lesions		
50	Female	Left pseudocyst
58	Female	Left polyp
34	Female	Left polyp
50	Female	Right polyp
27	Female	Nodules
61	Female	Nodules
19	Female	Prenodular swelling
55	Male	Left polyp

TABLE 2.
Descriptive Statistics for Measures of Phonation Quotient (PQ, in mL/s) in Each Disorder-Type Group and Instrument Type

Group	Instrument	Mean	Std. Deviation	*N*
Paralysis	Analog	567.38	228.03	9
	Spirometer	501.50	196.59	9
	PAS	550.27	223.84	9
Lesion	Analog	328.60	135.38	8
	Spirometer	309.69	124.08	8
	PAS	328.72	144.13	8
Total	Analog	455.02	221.61	17
	Spirometer	411.24	189.20	17
	PAS	446.01	217.10	17

Procedure

Participants began each session with three trials of MPT and then completed three trials of VC measurements on each of the instruments. The order of the instruments was counterbalanced. To complete the MPT tasks, participants were instructed to take a deep breath and sustain the vowel /a/ for as long as possible until they completely ran out of air. The investigator measured the duration using a stopwatch. The participants rested for up to 1 minute between trials. For the VC tasks, participants were provided the same instructions for all three instruments. Each participant was asked to take in a deep breath, place their mouth around the mouthpiece of the spirometer or in the facemask of the PAS ensuring a tight seal. The participant then exhaled maximally into the device. The participants wore a nose clip on the trials using a spirometer to prevent nasal air escape. The task of deep inhalation followed by maximum exhalation was demonstrated to the participants. If the participant began to exhale before placing the mouthpiece/facemask of the device, the participant was instructed to restart. The "Vital Capacity" protocol was used on the PAS. During this task, participants wore the PAS facemask with instructions to "press firmly" against the face, while the mask covered the mouth and nose.

Analyses

The largest VC trial (in milliliters) on each instrument was divided by the longest MPT trial (in seconds) to derive PQ measures. Statistical analyses were performed using *IBM SPSS* 23.0.[21] A univariate analysis of variance (ANOVA) was applied to PQ measures as a function of disorder type (paralysis vs lesion) and instrument (windmill-type, SP10, and PAS). For *post hoc* analysis of any significant main effects or interaction, paired samples *t* tests were used to compare PQ measurements. Correlational analyses were completed using a Pearson product-moment correlation.

RESULTS

Means and standard deviations for patients with paralysis/paresis or benign lesions on measures of PQ derived from the three different instruments are provided in Table 2. As a function of disorder type, measures of PQ were greatest in patients with paralysis/paresis, on average differing from patients with benign lesions by approximately 200 mL/s. There was a substantial degree of variability in each group data set. As a function of instrument type, measures of PQ were relatively similar, although those derived using the SP10 (digital spirometer) were always lowest.

A univariate two-way ANOVA applied to repeated measures of PQ with disorder group (paralysis/paresis vs benign lesion) and instrument type (windmill, SP10, PAS) was conducted. Results indicated a significant main effect of disorder group ($F[1,45] = 21.15$, $P < 0.001$) but no main effect for instrument type ($F[1,45] = 0.25$, $P = 0.77$). There was no significant interaction effect. These results indicated that patients with paralysis/paresis exhibited significantly greater PQ than patients with benign lesions, but instrument type did not have a significant effect on those measurements.

Pearson product-moment correlations were used to examine parallel forms reliability between instruments on measurements of PQ. Results of this analysis are provided in Table 3. All correlations were significant at the $P < 0.001$ level (two-tailed). Strong significant and positive correlations were observed on measures of PQ for all bivariate comparisons, with r^2 also large for each. These findings indicated that measures of PQ derived from each instrument type were strongly related to each other in a positive direction.

DISCUSSION

The purpose of this study was to measure parallel forms reliability on measures of PQ across three types of aerodynamic instruments representing low-tech/cost and high-tech/cost options for measuring VC, and to determine the characteristics of PQ in two disordered voice groups. The analog windmill-type and the digital SP10 spirometers were used as low-tech instrumentation, while the PAS served as the high-tech instrument for

TABLE 3.
Correlation Coefficients (*r*) and Coefficients of Determination (r^2) for Each Bivariate Pearson Correlation

		WindmillPQ	SP10PQ	PASPQ	
WindmillPQ	*r*			0.969**	0.938**
	r^2			0.94	0.88
	Sig. (two-tailed)			<0.001	<0.001
	N			17	17
SP10PQ	*r*	0.969**			0.950**
	r^2	0.94			0.90
	Sig. (two-tailed)	<0.001			<0.001
	N	17			17
PASPQ	*r*	0.938**	0.950**		
	r^2	0.88	0.90		
	Sig. (two-tailed)	<0.001	<0.001		
	N	17	17		

** Correlation is significant at the 0.01 level (two-tailed).

measuring VC. Previous research with the same instrumentation showed strong correlations between the three instruments, with the strongest correlation between the analog spirometer and the PAS.[8,9] Since disordered voice is associated with a decrease in MPT[5,15,22] and an increase in airflow rate from a glottal gap, it was important to determine if measures of PQ derived from different instruments reflected clinical expectations for conditions resulting in glottal insufficiency.

The PQ values obtained with the spirometers are in keeping with values previously reported in the literature for unilateral vocal fold paralysis and benign lesions.[5,17] The two disorder groups were significantly different on measures of PQ across instruments. This finding is consistent with previously published literature using measures of PQ and also measures of transglottal airflow derived from pneumotachograph systems such as the PAS. Existing reports have found that estimates of airflow using PQ or measures of transglottal airflow are substantially larger in patients with paralysis/paresis compared to those with benign lesions.[5,23] In the present study, across all three instruments the paralysis/paresis group manifested larger PQ measures by approximately 200 mL/s compared to the benign lesion group. This is likely explained by a larger degree of glottal insufficiency in the former group,[5] although glottal gap measurements would be needed to confirm that supposition in the present study sample. Mean PQ measurements in both patient groups of this study were also greater than those we obtained from our previous normative studies. We have previously reported normative adult male PQ values of 214.38 mL/s (analog windmill-type spirometer), 183.76 mL/s (digital SP10 spirometer), and 220 mL/s (PAS) and normative female PQ values of 140.46 mL/s (analog windmill-type spirometer), 124.81 mL/s (digital SP10 spirometer), and 146.86 mL/s (PAS).[8,9]

Correlational analyses revealed strong relationships between the three instruments for PQ indicating good parallel forms reliability. We used PQ measures derived from VC measures of the PAS for the purpose of measurement reliability from a high-tech pneumotachograph system. When combined with the nonsignificant ANOVA on measures of PQ using VC measures from the three instruments, it is reasonable to suggest that PQ derived from VC measures of low tech hand-held spirometers is a valid clinical option. This does not mean that one can directly compare measures of PQ to measures of transglottal airflow from high-tech pneumotachograph systems (eg, compare pretreatment PQ measures from one to post-treatment PQ measures from the other in the same patient). However, data from the present study do indicate that measurement reliability for PQ is strong when using VC measures from hand-held spirometers and high-tech systems used in this investigation. It is important for the clinician to use the same spirometer when comparing measurements within the same person, and also use the PQ data reported in the literature for comparison and not mean airflow rates.

Limitations

This study compared only two types of voice disorders. It is possible that with different pathologies, the data may not be as strong; however, since the effect of the disorder will be the same across instruments, one can expect PQ to be a valid option across disorder types. The groups were small and were not balanced for gender. There were three males in the paralysis/paresis group compared to one male in the benign lesion group. The study was limited to specific instrumentation and we cannot generalize the findings of this study to other instruments and the reader is recommended to use caution when using different instruments. Lastly, the spirometer does not provide us with information on subglottal pressure and laryngeal resistance, key aerodynamic measures.

CONCLUSIONS

PQ is a feasible assessment tool for the normal and disordered voice population as measured on the spirometers used in this study. The analog, digital, and pneumotachograph-based systems were reliable when deriving PQ from measurements of VC obtained from each. This supported a determination of good parallel forms reliability for PQ measures across the instruments. The PQ values derived from VC measurements of the spirometers were not significantly different from those derived from the PAS, further validating its use in clinical settings with limited resources. Clinicians using different spirometers from those used in this study should exercise caution and perform reliability checks prior to using them in the clinic. The data from our previous studies in addition to this data support the use of spirometers for VC and derived PQ measures in adults with normal and disordered voice.

REFERENCES

1. Stemple JC, Roy N, Klaben BG. *Clinical Voice Pathology: Theory and Management*. 5th ed. San Diego, CA: Plural Publishing; 2014.
2. Dejonckere PH. Assessment of voice and respiratory function. In: Remacle M, Eckel HE, eds. *Surgery of Larynx and Trachea*. Berlin, Heidelberg: Springer-Verlag; 2010.
3. Dejonckere P, Bradley P, Clemente P, et al. A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. *Eur Arch Otorhinolaryngol*. 2001;258:77–82.
4. Hirano M, Koike Y, von Leden H. Maximum phonation time and air usage during phonation. *Folia Phoniatr (Basel)*. 1968;20:185–201.
5. Iwata S, von Leden H. Phonation quotient in patients with laryngeal diseases. *Folia Phoniatr (Basel)*. 1970;22:117–128.
6. Rau D, Beckett RL. Aerodynamic assessment of vocal function using hand-held spirometers. *J Speech Hear Disord*. 1984;49:183–188.
7. Zraick RI, Smith-Olinde L, Shotts LL. Adult normative data for the KayPENTAX Phonatory Aerodynamic System Model 6600. *J Voice*. 2012;26:164–176.
8. Joshi A, Watts CR. Measurement reliability of phonation quotient derived from three aerodynamic instruments. *J Voice*. 2016;30:773.. e713-773.e719.
9. Joshi A, Watts CR. Phonation quotient in women: a measure of vocal efficiency using three aerodynamic instruments. *J Voice*. 2017;31:161–167.
10. Joshi A, Watts CR. A comparison of indirect and direct methods for estimating transglottal airflow rate. *J Voice* 2017; pii: S0892-1997(17) 30357-0 In press.
11. Choi J-N, Kim H, Nam CM, Choi H-S. Characteristics of glottal area waveform and phonation in patients with unilateral vocal fold palsy or sulcus vocalis. *Commun Sci Disord*. 2007;12:487–507.
12. Hertegård S, Hallén L, Laurent C, et al. Cross-linked hyaluronan versus collagen for injection treatment of glottal insufficiency: 2-year follow-up. *Acta Otolaryngol*. 2004;124:1208–1214.
13. Remacle M, Lawson G, Watelet J-B. Carbon dioxide laser microsurgery of benign vocal fold lesions: indications, techniques, and results in 251 patients. *Ann Otol Rhinol Laryngol*. 1999;108:156–164.
14. Remacle M, Mayné A, Lawson G, Jamart J. Subtotal carbon dioxide laser arytenoidectomy by endoscopic approach for treatment of bilateral cord immobility in adduction. *Ann Otol Rhinol Laryngol*. 1996;105:438–445.
15. Sjögren EV, van Rossum MA, Langeveld TM, et al. Voice outcome in t1a midcord glottic carcinoma: laser surgery vs radiotherapy. *Arch Otolaryngol Head Neck Surg*. 2008;134:965–972.
16. Verdonck-De Leeuw IM, Hilgers FJM, Keus RB, et al. Multidimensional assessment of voice characteristics after radiotherapy for early glottic cancer. *Laryngoscope*. 1999;109:241–248.
17. Wang C-C, Chang M-H, Jiang R-S, et al. Laryngeal electromyography-guided hyaluronic acid vocal fold injection for unilateral vocal fold paralysis: a prospective long-term follow-up outcome report. *JAMA Otolaryngol Head Neck Surg*. 2015;141:264–271.
18. Baken R, Orlikoff R. *Clinical Measurement of Speech and Voice*. 2nd ed. Clifton Park, NJ: Thompson Learning; 2000.
19. DeVon HA, Block ME, Moyle–Wright P, et al. A psychometric toolbox for testing validity and reliability. *J Nurs Scholarsh*. 2007;39:155–164.
20. KayPENTAX Corp LP, NJ. Instruction Manual, Phonatory Aerodynamic System (PAS), Model 6600.
21. IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp; 2013.
22. Yumoto E. Aerodynamics, voice quality, and laryngeal image analysis of normal and pathologic voices. *Curr Opin Otolaryngol Head Neck Surg*. 2004;12:166–173.
23. Dastolfo C, Gartner-Schmidt J, Yu L, Carnes O, Gillespie AI. Aerodynamic outcomes of four common voice disorders: moving toward disorder-specific assessment. *J Voice*. 2016;30:301–307.