Using a Vortex Whistle System to Estimate Phonatory Airflow via the Phonation Quotient

[#,†]Shaheen N. Awan, [‡]Jordan A. Awan, [§]Heather Shaw Bonilha, [¶]Amanda Gillespie, [#,†]Victoria McKenna, [‖]Jun Chen, and [#,†]David A. Eddins, #†*Orlando, Florida,* ‡‖*West Lafayette, Indiana,* §*Columbia, South Carolina, and* ¶*Atlanta, Georgia*

SUMMARY: Introduction. The primary purpose of this study was to examine the relationship between phonation quotient (PQ) obtained using a low-cost vortex whistle system (VWS) vs phonatory airflow (PA) in sustained voicing calculated using a clinical-standard pneumotach-based system (the Phonatory Aerodynamic System [PAS]). PQ values obtained using the VWS were also compared with PQ values obtained using a "gold" standard pneumotach-based Koko Sx1000 spirometer.

Methodology. Participants were 91 vocally healthy young adults (61 females and 30 males) between the ages of 18 and 30 years. Measures of vital capacity (VC) were obtained using the VWS and Koko Sx1000 spirometer, and measures of maximum phonation time (MPT) and PA during sustained phonation were obtained from the Maximum Sustained Phonation protocol on the PAS. PQs were calculated via PQ = VC (L)/MPT (s).

Results. Correlations of r = 0.828 and r = 0.791 were observed between PA and PQ_{Koko} and PQ_{VWS}, respectively, with no significant difference between the two correlations. Stronger correlations between both PQ_{Koko} and PQ_{VWS} and PA were observed in males vs females. Linear regressions to predict PA from either PQ_{Koko} or PQ_{VWS} resulted in residual standard errors of 0.037 L/s (37 mL/s) and 0.040 L/s (40 mL/s), respectively. PQ_{VWS} and PQ_{Koko} were highly correlated (r = 0.961).

Conclusion. These results confirm prior research demonstrating that PQ may be used as a reasonable substitute for PA, with error values within previously reported expected ranges of intrasubject variability. In addition, a low-cost VWS can provide similarly accurate measures of PQ vs those computed using VC estimates from a "gold standard" pneumotach-based spirometer. PQ estimates of airflow via a VWS show great promise as a low-cost method that can promote the use of respiratory measures of capacity and flow to the voice assessment protocols of speech-language pathologists.

Key Words: Phonatory airflow—Phonation quotient—Vortex whistle—Vital capacity—Maximum phonation time.

INTRODUCTION

Aerodynamic measures are one of the five domains that comprise the multidimensional voice evaluation as supported by multiple clinical practice guidelines.[1–3] Phonatory airflow (PA) (also known as mean flow rate; glottal airflow; transglottal airflow; phonatory flow rate; mean expiratory airflow; average airflow) has been used for years in the clinical voice laboratory. PA (the amount of air moving between the vocal folds during phonation; the average volume of air flowing through the glottis over a specific period of time, measured in L/s or mL/s) is a valuable index of vocal function. Because PA is influenced by factors related to phonatory physiology such as respiratory drive, glottal configuration, medial compression force, and tissue stiffness,[4] PA can provide valuable information in the description of typical vs dysphonic voice. Speakers with dysphonia associated with glottal incompetence (eg, secondary to conditions such as vocal fold paralysis, vocal fold bowing, or the presence of mid-membranous lesions) or problems with respiratory-PA management tend to exhibit higher airflow rates than speakers with typical voice production.[5–11] In contrast, low flow rates may be indicative of conditions such as spasmodic dysphonia or behavioral dysphonias such as hyperadducted muscle tension dysphonia.[10,12–14] Thus, evaluating flow rate is important to understand the nature of the voice problem. This measure can also be used to indicate treatment success,[12] with voice therapy or surgery that improves glottal closure resulting in a decrease in PA[5,10,13,14] and therapies that reduce laryngeal hyperfunction tending to result in an increase in measures of airflow.[10,13]

Unfortunately, regardless of the documented value of PA measures and despite best-practices recommendations,[2,15,16] surveys of speech-language pathology (SLP) assessment and treatment practices in the United States[17,18] indicate that aerodynamic measures such as the

Accepted for publication August 5, 2025.

Research reported in this publication was supported by the National Institute on Deafness and Other Communication Disorders of the National Institutes of Health under Award Number R01DC020799 (Awan SN, Awan JA, Chen J, and Gillespie A). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

From the *Communication Technologies Research Center, University of Central Florida, Orlando, Florida; †School of Communication Sciences and Disorders, University of Central Florida, Orlando, Florida; ‡Deptartment of Statistics, Purdue University, West Lafayette, Indiana; §Department of Communication Sciences and Disorders, University of South Carolina, Columbia, South Carolina; ¶Emory University School of Medicine, Atlanta, Georgia; and the ‖School of Mechanical Engineering, Purdue University, West Lafayette, Indiana.

Address correspondence and reprint requests to Shaheen N. Awan, Communication Technologies Research Center, University of Central Florida, 3280 Progress Drive, Suite 100, Orlando, FL 32826. E-mail: shaheen.awan@ucf.edu

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

0892-1997

https://doi.org/10.1016/j.jvoice.2025.08.008

measurement of PA are rarely used, a clinician compliance issue that also exists internationally.[19–22] Key factors that limit the use of aerodynamic measures in clinical practice are (a) cost and (b) availability of clinicians with training and experience. PA is typically measured using a pneumotach-based system. A pneumotach (aka, pneumotachograph; pneumotachometer) is an instrument used to measure respiratory flow via a differential pressure transducer in which flow is directly proportional to the pressure drop across a resistive element. However, systems that incorporate pneumotachs for voice analyses (eg, Phonatory Aerodynamic System—PAS, Pentax Medical Inc., Montvale, NJ; Aeroview, Glottal Enterprises Inc., Syracuse, NY) can cost upward of several thousands of dollars. In turn, this relatively high cost restricts the availability of this type of instrumentation in many SLP training programs, limiting clinician familiarity with the acquisition and use of measures such as PA. In turn, outside of focused voice centers, these costs reduce the availability of this instrumentation in clinical practice.

A potential low-cost alternative to pneumotach-measured PA is the phonation quotient (PQ), where PQ = Vital Capacity [VC, L]/Maximum Phonation Time [MPT, s]. PQ has been described as a "reasonable clinical substitute" (p. 199)[23] for measurement of airflow rate obtained during phonation when PA cannot be directly determined via instrumentation such as a pneumotach.[23,24] Hirano et al[23] investigated PA during phonation in 50 normal voice and 73 organic and functional voice-disordered participants and reported significant correlations of r's = 0.86 and 0.75 in normophonic males and females (respectively) between PQ and PA obtained via a pneumotachograph during maximally sustained phonation. These authors observed that PA and PQ observations aligned with similar clinical categories (< normal; normal; > normal) in 86% of their pathological cases and concluded that the PQ may serve as a fairly accurate clinical substitute for the PA. Iwata and von Leden[24] examined the relationship between PQ and airflow (via pneumotachograph) in maximally sustained phonation in 192 patients with laryngeal diseases and reported correlations between PA and PQ as r = 0.76 in chronic laryngitis, r = 0.89 in benign tumors, r = 0.62 in malignant tumors, and r = 0.90 in cases of unilateral vocal fold paralysis. Yoshioka et al[25] compared PQ and PA via spirometer in 50 normophonic adults and 180 pathological cases and obtained PA in two conditions: maximum sustained phonation after maximum inspiration and comfortable sustained phonation (3 to 5 seconds) after typical breathing. These authors reported significant and strong correlations between PQ and PA during maximum sustained phonation, with stronger results observed in males (r = 0.88) vs females (r = 0.68). Though still significant, weaker correlations between PQ and PA from comfortable sustained phonation were observed, with the result attributed to increased sample-to-sample variability related to the ambiguity of effort required during "normal/comfortable" voice production. Rau and Beckett[26] reported a high, positive correlation between actual airflow rates and estimated PA via the PQ using three different portable spirometers (r's ranged from 0.74 to 0.87, $P < 0.0001$) in 19 participants. In addition, these authors combined their data with data obtained via similar methodology by Yanagihara et al[27] to produce a regression equation for adjusting measures of PQ to provide a more accurate representation of PA. This adjusted measure was referred to as estimated mean flow rate (eMFR) and was calculated as follows: eMFR = 77 + 0.236 × PQ. In addition to these aforementioned early studies investigating the validity of the PQ, other studies have examined the reliability of obtaining PQ with different pneumotach vs portable spirometers,[28] comparison of PA measured via pneumotach vs PQ,[29] and have used the PQ measure in the evaluation of both typical and disordered voice.[30–34]

However, the ability to calculate PQ is itself limited by its dependence on an accessible and accurate measurement of VC. Measures of VC typically require the use of a spirometer, costing from hundreds of dollars for portable handheld models to thousands of dollars for high-end pneumotachometer systems. However, the vortex whistle (VW), a low-cost device that produces an acoustic signal with a frequency directly proportional to the airflow moving through it, [35,36] has shown strong potential for providing respiratory measures such as VC.[37] Once the acoustic output of the VW is captured, analysis algorithms can map signal frequency to the inlet airflow rate (in L/s) and the area under the frequency (ie, flow) curve can be used to calculate expiratory volume (in L). Recent studies by Awan and Awan[38,39] have validated the accuracy of VC measures obtained via a vortex whistle system (VWS) consisting of a 3-D printed VW and custom analysis software.

The primary purpose of this study was to examine the relationship between PQ obtained using a low-cost VWS vs the PA calculated using a clinical-standard pneumotach-based system (the PAS) during sustained voicing. In accord with previous studies,[23–26] it was hypothesized that PQ obtained using a VWS would correlate strongly with PA obtained via a pneumotach-based device. A secondary aim was to compare PQ values obtained using the VWS to PQ values obtained using a "gold standard" pneumotach-based spirometry system (the Koko Sx1000 spirometer, nSpire Health). Since the VWS has been demonstrated to provide an accurate measurement of VC, with a strong correlation to the VC obtained by the Koko spirometer, it was hypothesized that PQ measures obtained using the VWS would be highly correlated with PQ measures obtained using the Koko Sx1000 spirometer.

METHODOLOGY

The data collection methodology was approved by the Bloomsburg University Institutional Review Board (No. 2018-118) and has been described in detail in Awan and Awan.[38] Participants were 91 healthy young adults (61

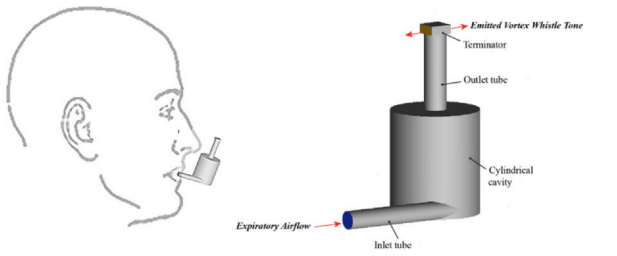

FIGURE 1. Positioning of the vortex whistle during vital capacity testing (left) and component sections of the vortex whistle (right).

females and 30 males) between the ages of 18 and 30 years with typical voice as evaluated by a speech-language pathologist with expertise in the assessment of voice disorders (author S.N.A.). Participants were nonsmokers, reported no significant history of untreated respiratory disorder, no report of cough/cold or seasonal allergies that may have affected breathing on the day of testing, and reported no significant history of known neurological disorder.

Participants were tested using (1) the PAS (Model 6600, Pentax Medical Inc., Montville, NJ), (2) the Koko Sx1000 spirometer, and (3) the VWS, comprising a 3-D printed VW and audio recording and analysis software.[38] The VWs were manufactured using a Flashforge Creator Pro 3-D printer (Flashforge USA, City of Industry, CA) via the open-source slicing software *FlashPrint* using nontoxic, biodegradable polylactic acid plastic. The VW design dimensions used in this study have been described in Awan and Awan.[38] Figure 1 provides views of the VW in use and the components of the VW.

Experimental tasks

Participants were asked to complete the following:

1. VC using a pneumotach-based spirometer and the VWS: For the purposes of this study, VC was elicited via the forced vital capacity (FVC—exhale with maximal force after deepest possible inhalation). While many PQ studies do not specify how VC was elicited, the FVC definition is consistent with the elicitation description provided by Rau and Beckett.[26] Task instructions were consistent with those provided for the KoKo Sx1000 spirometer and task performance was supported via encouragement of the examiner:
   Following 2 to 4 normal breaths; "On your next breath, take a deep fast breath"; "Place lips firmly around the inlet tube and blast it out hard and fast"; "Keep blowing, keep blowing, keep blowing all the way out until you are empty."
   The order of measurement system (KoKo vs VWS) was counterbalanced across participants to account for any potential order effect. For the Koko Sx1000 spirometer, all FVC elicitations were produced with the subject wearing a nose-clip. The KoKo Sx1000 spirometer was connected via USB to a PC laptop running Windows 7 and FVC productions were elicited using a disposable filter. Three trials were elicited and recorded using the *Koko Sx1000* software with a 60-second rest period between trials. Measurements of FVC were automatically computed using the *Koko Sx1000* software.
   Three trials of FVC via the VW were also obtained with all participants wearing a nose-clip during the task. VWs were sterilized externally and internally prior to subject use. During FVC elicitation, three trials were elicited with a 60-second rest period between trials. The VW signal was recorded for each trial using a Shure Beta-54 headset microphone (Shure Inc.) with a microphone-to-VW outlet tube distance of ≈ 25 mm. The headset microphone was connected to the microphone input of a Yamaha Audiogram 6 digital interface that was connected to a PC laptop via USB. The VW tone was recorded at 44.1 kHz, 16 bits and saved in .wav format using custom *Java* software. VC from the VWS was calculated using the same custom *Java* software and adjusted using a frequency-to-flow conversion formula followed by a VC regression formula reported in Awan and Awan[38] that transforms the raw area under the flow curve to approximate the VC value provided via the gold standard Koko Sx1000 spirometer.
2. The Maximum Sustained Phonation protocol of the PAS: While holding the PAS device and mask firmly over the nose and mouth, the subject was instructed to take a deep breath and sustain the vowel /a/ at a comfortable pitch and loudness for as long as possible. Three trials were elicited. Prior to use, the PAS facemask was sterilized and the PAS flowhead was calibrated according to manufacturer's instructions. Three trials were elicited, with the longest duration production selected as the MPT (seconds) and the PA (L/s) (PAS terminology: mean expiratory flow rate) obtained from the PAS statistics display.
3. PQ Calculation: For each subject, the PQ was calculated using the largest FVC (from the Koko system trials and the VWS trials, respectively) and the longest MPT from the PAS MPT protocol trials. To maintain consistency with previous descriptions of the PQ,[23,24,26] "VC" will be used in place of "FVC" in PQ formulas and further description of the method. The formulas for the PQ calculations were as follows:

$$PQ_{Koko} = \frac{VC(L)_{Koko}}{MPT(s)}$$

$$PQ_{VWS} = \frac{VC(L)_{VWS}}{MPT(s)}$$

Statistical analyses

Correlational and linear regression analyses were computed using *R* v.4.4.1 statistical computing software[40] and

conducted to ascertain the presence and strength of the relationships between (a) PQ measurements using the VWS and Koko Sx1000 vs PA and (b) PQ measurements obtained using the VC from the Koko Sx1000 vs the VWS. Correlation coefficients range from 0 to ± 1, with values from ± 0.1 to 0.3 considered weak; ± 0.4 to 0.7, moderate; and ± 0.8 to 0.99, strong to very strong.[41] Holm-Bonferroni corrections for multiple correlations were used to adjust the family-wise error rate. Comparison of correlations from dependent groups was assessed using Steiger's z,[42] and correlations from independent groups were assessed using Fisher's z.[43] Correlation comparisons were computed using the *cocor: comparing correlations* website.[44,45] In all linear regression models considered, diagnostic plots demonstrated that the model assumptions were reasonable.

A two-way Analysis of Variance (ANOVA; 3 levels of Method of measuring airflow [PA, PQ_{Koko}, and PQ_{VWS}]; 2 levels of Sex) was computed using *JASP v.0.95.1*[46] to investigate potential differences in PA among methods and between sexes. In the event of violations of sphericity, ANOVA results were evaluated using Greenhouse-Geisser corrections. For each ANOVA, effect sizes were computed using eta^2 (η^2), with a small effect ≥0.01, a moderate effect ≥0.06, and a strong effect ≥0.14.[47] ANOVA results with negligible effects (ie, < small effect sizes) are not discussed. Post hoc analyses of significant ANOVAs were evaluated with *t* tests and Holm-Bonferroni corrections for family-wise error. Post hoc effect sizes were evaluated using Cohen's *d* (small [0.2], medium [0.5], and large [0.8] effects).[47]

RESULTS

PQ vs PA

The results of a two-way ANOVA (3 levels of Method of measuring PA [PA; PQ_{Koko}; PQ_{VWS}]; 2 levels of Sex) showed a significant main effect of Method with a small effect size (Greenhouse-Geisser: $F = 22.332$, $P < 0.001$, $\eta^2 = 0.022$). The main effect of Sex was nonsignificant with a small effect size ($F = 1.890$, $P = 0.173$, $\eta^2 = 0.019$), and the interaction of Method × Sex was also nonsignificant with a negligible effect size (Greenhouse-Geisser: $F = 2.238$, $P = 0.129$, $\eta^2 = 0.002$).

Post hoc *t* test analysis of the main effect of Method showed significant small effect size differences between PA and PQ_{Koko} (mean difference = −0.025 L/s, $t = -5.16$, $P < 0.001$, Cohen's $d = -0.33$) as well as between PA and PQ_{VWS} (mean difference = −0.027 L/s, $t = -4.86$, $P < 0.001$, Cohen's $d = -0.35$). No significant difference was observed between PQ_{VWS} and PQ_{Koko} (mean difference = −0.002 L/s, $t = -0.531$, $P = 0.60$, Cohen's $d = -0.018$). Table 1 provides summary values from the different Methods and Table 2 for summaries broken down by Sex. Figure 2 provides combined scatterplots, box plots, and distribution plots for PQ_{Koko} vs PA and PQ_{VWS} vs PA, respectively.

The following analyses establish how well either PQ measure correlates with PA via the PAS. PQ_{Koko} and PA had a strong correlation (r = 0.828; 95% confidence intervals [CI] = 0.751, 0.883; $r^2 = 0.686$), while PQ_{VWS} and PA had a high moderate-to-strong correlation (r = 0.791; 95% CI = 0.699, 0.857; $r^2 = 0.626$). A comparison of two overlapping correlations based on dependent groups was used to compare the correlations between PQ_{Koko} and PA vs PQ_{VWS} and PA, while also taking into account the correlation between PQ_{VWS} vs PQ_{Koko} (see below). Using a Holm-Bonferroni correction for family-wise error (*P* values are multiplied by 3), the results indicated no significant difference between the two correlations (Steiger's z = 2.14, $P = 0.09$). In addition, substantial overlap between the respective correlation 95% CI also confirmed the two correlations as being not significantly different.

Linear regressions to predict PA from either PQ_{Koko} or PQ_{VWS} resulted in the following conversion formulas:

$$PA = -0.004 + (0.893 \times PQ_{Koko}),$$

$$PA = 0.015 + (0.787 \times PQ_{VWS}),$$

with residual standard errors of 0.037 L/s (95% CI = 0.032, 0.043) and 0.040 L/s (95% CI = 0.036, 0.047), respectively. In both models, the coefficient of PQ was significant at $P < 0.001$. As these intervals overlap, the regression errors between the two models are not significantly different. Plots of PQ_{Koko} vs PA and PQ_{VW} vs PA are shown in Figure 3.

Regressions to predict PA from either PQ_{Koko} or PQ_{VWS} were also considered with an additional main effect of Sex. While the effect was statistically significant ($P = 0.0312$ for Koko and $P = 0.0355$ for VWS), comparison of the Akaike

TABLE 1.
Phonatory Airflow (PA, L/s) Measured Using the PAS vs PQ-Estimated PA Using Vital Capacity (VC) Measured Using Either the Koko Sx1000 Spirometer or the Vortex Whistle System (VWS) Across All Participants (N = 91)

Statistic	PA via PAS (L/s)	MPT (s)	VC_{Koko} (L/s)	VC_{VWS} (L/s)	PQ_{Koko} (L/s)	PQ_{VWS} (L/S)
Mean	0.171	24.02	4.435	4.397	0.196	0.198
Standard deviation	0.065	7.58	0.90	0.89	0.061	0.066
Minimum	0.067	12.48	3.05	3.02	0.096	0.100
Maximum	0.460	44.17	7.15	7.05	0.378	0.437

Abbreviations: L/s, liters per second; MPT, maximum phonation time (s); PAS, Phonatory Aerodynamic System; PQ_{Koko}, phonation quotient using the Koko vital capacity estimate; PQ_{VWS}, phonation quotient using the VWS vital capacity estimate.
Notes: MPT (s) and VCs from the Koko Sx1000 and the VWS are also provided.

TABLE 2.
Phonatory Airflow (PA, L/s) Measured Using the PAS vs PQ-Estimated PA Using Vital Capacity (VC) Measured Using Either the Koko Sx1000 Spirometer or the Vortex Whistle System (VWS) for Males (n = 30) vs Females (n = 61)

Sex	Statistic	PA via PAS (L/s)	MPT (s)	VC_{Koko} (L/s)	VC_{VWS} (L/s)	PQ_{Koko} (L/s)	PQ_{VWS} (L/s)
Males	Mean	0.191	28.53	5.38	5.32	0.204	0.208
	Standard deviation	0.078	8.51	0.79	0.78	0.072	0.082
	Minimum	0.080	12.48	4.11	4.15	0.096	0.102
	Maximum	0.460	44.17	7.15	7.05	0.378	0.437
Females	Mean	0.162	21.83	3.98	3.94	0.193	0.193
	Standard deviation	0.056	6.03	0.52	0.51	0.055	0.057
	Minimum	0.067	12.56	3.05	3.02	0.103	0.100
	Maximum	0.323	34.64	5.35	5.29	0.326	0.326

Abbreviations: L/s, liters per second; MPT, maximum phonation time (s); PAS, Phonatory Aerodynamic System; PQ_{Koko}, phonation quotient using the Koko vital capacity estimate; PQ_{VWS}, phonation quotient using the VWS vital capacity estimate; VC: vital capacity.
Notes: MPT (s) and VCs from the Koko Sx1000 and the VWS are also provided.

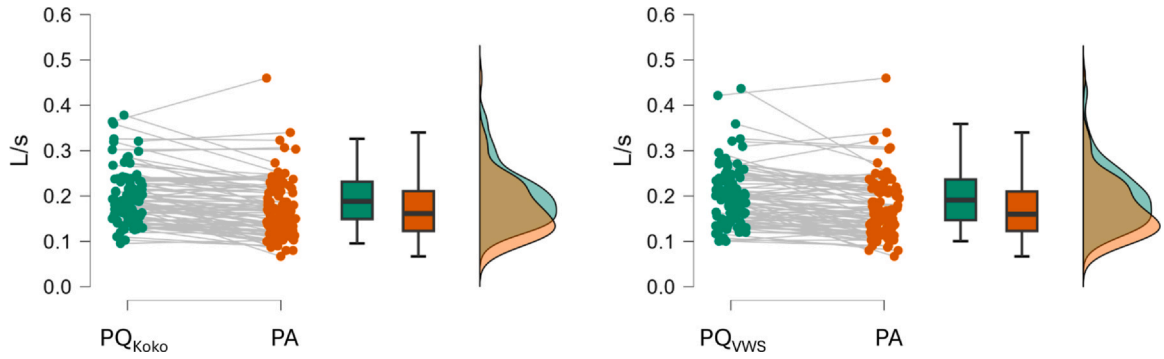

FIGURE 2. Combined scatterplots, box plots, and distribution plots for PQ_{Koko} vs PA (left) and PQ_{VWS} vs PA (right). PA, phonatory airflow; PQ_{Koko}, phonation quotient using the Koko vital capacity estimate; PQ_{VWS}, phonation quotient using the vortex whistle system vital capacity estimate.

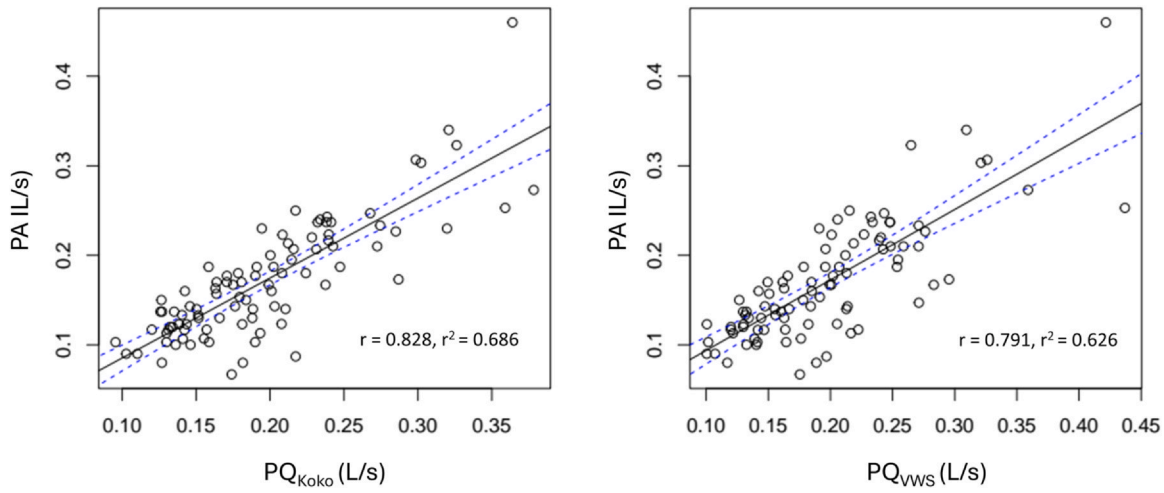

FIGURE 3. Plots of PQ_{Koko} against PA (left) and PQ_{VWS} against PA (right), along with regression lines and 95% confidence bands. PA, phonatory airflow; PQ_{Koko}, phonation quotient using the Koko vital capacity estimate; PQ_{VWS}, phonation quotient using the vortex whistle system vital capacity estimate.

information criterion (AIC) revealed differences of only 2.8 for Koko and 2.6 for VWS. According to Hilbe,[48] because the sample size is ≤256 and the difference in AIC is < 6, the simpler model that does not include Sex as a predictor is preferred.

Though the main effect of Sex was not strong enough to warrant including it in the model, separate correlations between PA and PQ were computed for males vs females since it is well-established that adult males and females can differ substantially in both VC and MPT components used in PQ calculations. Strong significant correlations were observed in males between PA and both PQ_{Koko} (r = 0.848; 95% CI = 0.702, 0.905; r^2 = 0.717) and PQ_{VWS} (r = 0.816; 95% CI = 0.645, 0.909; r^2 = 0.666), while significant but slightly weaker correlations between PA and PQ_{Koko} (r = 0.817; 95% CI = 0.712, 0.886; r^2 = 0.667) and PQ_{VWS} (r = 0.769; 95% CI = 0.641, 0.845; r^2 = 0.591) were observed in females. Comparison of correlations based on independent groups indicated no significant difference ($P \geq 1.0$ after correction) between males vs females for PA vs PQ_{Koko} correlations (Fisher's z = 0.435) or for PA vs PQ_{VWS} correlations (Fisher's z = 0.544).

PQ_{VWS} vs PQ_{Koko}

The correlation between PQ_{VWS} and PQ_{Koko} was strong (r = 0.961; 95% CI = 0.941, 0.974; r^2=0.923). A linear regression using PQ_{VWS} to predict PQ_{Koko} resulted in the following predictive formula:

$$PQ_{Koko} = 0.022 + (0.879 \times PQ_{VWS}),$$

with residual standard error of 0.017 L/s (95% CI = 0.0148, 0.0200). PQ_{VWS} was a highly significant predictor with $P < 0.001$. Thus, given PQ_{VWS}, the value of PQ_{Koko} can typically be predicted with error ≤0.02 (L/s). A plot of PQ_{VWS} vs PQ_{Koko} along with the regression line and 95% confidence band is shown in Figure 4. When Sex was included as a predictor, it was found to be nonsignificant (P = 0.578) and so was excluded from the model.

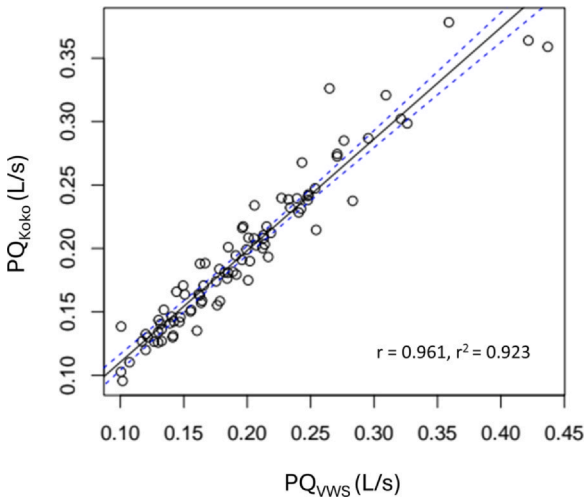

FIGURE 4. Plot of PQ_{VWS} vs PQ_{Koko} along with regression line and 95% confidence band. PQ_{Koko}, phonation quotient using the Koko vital capacity estimate; PQ_{VWS}, phonation quotient using the vortex whistle system vital capacity estimate.

DISCUSSION

Measures of PA can provide important information regarding the underlying physiology of voice production but has been limited in clinical use by the cost and availability of the instrumentation necessary to acquire airflow measurement. The results of this study are consistent with prior research that has established that the PQ can be a suitable approximation to the PA. Furthermore, the VWS can accurately estimate VC that can then be used to accurately estimate PQ. Because the potential cost of the VW and analysis software is a small fraction of the cost of a pneumotach-based airflow or spirometry system (as well as a fraction of the cost of portable spirometers), the VWS has the capability to expand the use of aerodynamic measures in clinical voice assessment by providing the VC data necessary for PQ estimates of airflow.

The strength of correlation between PA and PQ_{VWS} (r = 0.791) observed in this study is consistent with previous reports[23,24,26] and was similar to the strength of correlation between PA and PQ_{Koko} (r = 0.828). When adjusted to spirometry standards, airflow rates estimated via both PQ_{VWS} and PQ_{Koko} tend to overestimate pneumotach-captured airflow (0.198 and 0.196 L/s, respectively, vs 0.171 L/s [PA]). This is expected since the entire VC is generally not used during phonation.[26,49] The difference between phonation volume and VC, as well as differences in expiratory air pressures and flow rates during voiced vs nonvoiced tasks explain why PQ estimates typically overestimate actual PA.[24,26] The linear regression equations for the estimation of PA using VWS-derived PQ estimates confirm the general overestimation of PA by PQ by incorporating slope coefficients that are < 1.0, thereby reducing PQ estimates to better approximate actual PA. While the same form for the regression equation for predicting PA was reported by Rau and Beckett[26] (referred to as eMFR), the y-intercept and slope coefficients are quite different than reported here for the VWS due to factors such as substantial differences in the instrumentation used to estimate PQ and the sample size (37 participants in Rau and Beckett[26] vs 91 participants in the current study).

The residual standard errors in the regressions for PQ_{VWS} or PQ_{Koko} vs PA were observed as 0.037 L/s (37 mL/s) and 0.040 L/s (40 mL/s), respectively. The potential accuracy of the PQ is reinforced by the fact that these errors (21.64% and 23.39%, respectively; calculated by residual standard errors divided by the mean PA [0.171 L/s]) are consistent with previously reported expected ranges of variability within the same person. In a study examining aerodynamic and electroglottographic reliability via coefficients of variation (CVs) in 21 subjects (11 males and 10 females) recorded on 4 nonconsecutive days over a 2-week

period for each subject, Higgins et al[50] reported that, for syllable repetition, average day-to-day intrasubject variation for mean PA was approximately 15%, representing a day-to-day change of approximately 20 to 30 mL/s. These authors further stated that > 25% variation may be necessary to reflect "a real change in phonatory status" (p. 42). Awan et al[51] examined intrasubject variability in aerodynamic measures of voice in 60 subjects (30 males and 30 females) on 2 separate days with approximately 1 week between each session at approximately the same time of day. These authors reported typical standard error of measurement for PA as 25 to 35 mL/s (associated CV = 18 to 20%) and a minimum difference in PA (ie, an estimate of the difference needed between separate measures on a subject to be considered real) as conservatively measured between 53 and 77 mL/s. Therefore, the results of this study show that PQ not only correlates well with PA but also provides an estimate of PA that, in this study, is well within expected day-to-day variability.

Comparison of PQs using different methods for eliciting VC

The correlation between PQ_{VWS} and PQ_{Koko} was very strong (r = 0.961, r^2 = 0.923), indicating that PQ calculations using VC estimates from "gold standard" pneumotach-based spirometry systems can be replicated with high precision via lower-cost alternatives such as the VWS used in this study. This finding is consistent with previous studies by Rau and Beckett[26] and Joshi and Watts[28,30] who also reported strong correlations between PQs calculated using VC measures from various forms of instrumentation (eg, wet respirometer; portable handheld turbine spirometers; mass flow transducers; ultrasonic sensors; pneumotach-based methods). This result indicates that beneficial aerodynamic measures related to vocal function such as measures of respiratory volume and estimates of PA can be added to the voice assessment protocols of speech-language pathologists and other voice professionals with minimal costs.

PQs in males vs females

The mean PQs for male and female participants (Table 2) correspond well with the 0.10 to 0.20 L/s (100 to 200 mL/s) general expectation for typical airflow during vowel production[4] and are well within the critical regions for typical PQ production of 0.069 to 0.307 L/s for males and 0.078 to 0.241 L/s for females reported by Hirano et al.[23] In addition, the observation of higher airflow in males vs females (Table 2), regardless of method, is consistent with expectations of higher airflow in males due to their larger cross-sectional area of the glottis and lower airflow resistance in comparison with females.[4,52,53] The observation that the correlation between PA and PQ was stronger in males than females in this study has also been reported by Hirano et al[23] who observed correlations between PQ and mean flow rate of r = 0.85 in males vs r = 0.75 in females. Similarly, Yoshioka et al[25] reported stronger correlations between PQ and PA in males (r = 0.88) vs females (r = 0.68). It is unclear as to why PQ estimates should differ in males vs females. Table 2 shows that the mean difference between PA and PQ is approximately −0.14 to −0.17 L/s in males, while the mean PA-PQ difference is −0.30 to −0.32 L/s in females. A review of our data showed that approximately 13 to 15% of the participants in the female group produced particularly low mean PAs on the PAS vs their estimated PA via PQ calculations via the Koko or VWS (less than the mean PA-PQ difference minus 1 standard deviation; < −0.70 to −0.64 L). In contrast, 10% of the male participants showed particularly low mean PA's on the PAS (less than the mean PA-PQ difference minus 1 standard deviation; < −0.66 to −0.60 L). It may be that possible sex differences in phonatory vs nonphonatory volumes and/or behavior during maximum performance phonatory vs nonphonatory tasks may affect the strength of correlation between PA and PQ in males vs females, but further study is necessary.

Limitations

Because PQ is calculated using two maximum performance tests (VC and MPT), it is essential that the examiner monitors and encourages the performance of the subject to provide the best possible performance. Even with best practice, some participants may underachieve during a task such as the FVC or MPT. Also, in otherwise healthy participants, tasks such as the FVC may induce bronchospasm and fatigue, which may affect the repeatability of the task during multiple trials and may also affect the accuracy of the VC measurement. We have also observed that some participants will terminate the MPT once they feel that voice becomes effortful, again resulting in underestimations. While participants were monitored closely during the aforementioned tasks, it is possible that variations in maximum performance may have influenced the strength of reported correlation between PA and PQ. Future studies may incorporate the use of visual feedback to encourage maximum performance and decrease intertrial variability on VC and MPT tasks. In addition, alternative methods of eliciting VC (eg, slow vital capacity; aka, expiratory VC, extended VC—a slow maximal exhalation produced after the deepest possible inspiration), which are less fatiguing, which can also provide valid measures of respiratory capacity[39] and may aid in reliably eliciting the component measures that are used in the PQ calculation, should also be considered. Finally, while the effect of Sex was considered in the analyses, it would be preferred to have an equal representation of both male and female participants to assess possible sex differences in strength of correlation between PA and PQ.

CONCLUSIONS

The results of this study have reproduced prior research demonstrating the relationship between PQ and PA using a low-cost VWS and have also shown that the VWS can

provide similarly accurate measures of PQ when compared with PQ measures computed using VC estimates from a "gold standard" pneumotach-based spirometry system (Koko Sx1000). Future studies are planned to evaluate the VWS and measures of both VC and PQ across the lifespan and in persons with voice disorders, and to evaluate possible modifications of the VW and analysis software that may improve the accuracy of measures.

Currently, speech-language pathologists in the United States bill Current Procedural Terminology (CPT) code 92520 "Laryngeal function studies (i.e., aerodynamic testing and acoustic testing)" for reimbursement for objective testing during voice evaluation. However, the SLP must use modifier 52 (reduced service) if only one test is performed (ie, aerodynamic testing only, acoustic testing only).[54] Since many clinicians do not have access to an aerodynamics testing system such as the PAS, reimbursements for this CPT code are generally reduced since only acoustic testing is conducted. The ability to use a VWS for objective measures of respiratory capacity and PA via the PQ will allow clinicians to not only provide more comprehensive voice assessment but will make voice assessment more cost-effective by providing a low-cost tool that allows the clinician to bill CPT 92520 without the modifier and receive full reimbursement. The VW is a 3-D printed device that contains no moving parts, does not require an electrical or battery power source, and can be produced for a small fraction of the cost of both pneumotach-based systems and portable spirometers. Obtaining PQ estimates of airflow via a VWS shows great promise as a method that can increase the use of basic respiratory measures of capacity and flow in voice assessment protocols of speech-language pathologists and allow for improved compliance with clinical practice guidelines.

Data availability

Data supporting the results reported in this article may be obtained upon request from the contact author.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Lechien JR, Geneid A, Bohlender JE, et al. Consensus for voice quality assessment in clinical practice: guidelines of the European Laryngological Society and Union of the European Phoniatricians. *Eur Arch Otorhinolaryngol.* 2023;280:5459–5473. https://doi.org/10.1007/S00405-023-08211-6.
2. Patel RR, Awan S, Barkmeier-Kraemer J, et al. Recommended protocols for instrumental assessment of voice: American Speech-Language-Hearing Association Expert panel to develop a protocol for instrumental assessment of vocal function. *Am J Speech Lang Pathol.* 2018;27:887–905.
3. Stachler RJ, Francis DO, Schwartz SR, et al. Clinical practice guideline: hoarseness (dysphonia) (update). *Otolaryngol Head Neck Surg.* 2018;158(suppl 1):S1–S42. https://doi.org/10.1177/0194599817751030.
4. Watts C, Awan S. *Laryngeal Function and Voice Disorders: Basic Science to Clinical Practice*. New York, NY: Thieme Publishers; 2019.
5. Hartl DM, Hans S, Vaissière J, et al. Laryngeal aerodynamics after vocal fold augmentation with autologous fat vs thyroplasty in the same patient. *Arch Otolaryngol Head Neck Surg.* 2005;131:696–700. https://doi.org/10.1001/ARCHOTOL.131.8.696.
6. Cantarella G, Baracca G, Pignataro L, et al. Assessment of dysphonia due to benign vocal fold lesions by acoustic and aerodynamic indices: a multivariate analysis. *Logop Phoniatr Vocol.* 2011;36:21–27. https://doi.org/10.3109/14015439.2010.528022.
7. Ma EPM, Yiu EML. Multiparametric evaluation of dysphonic severity. *J Voice.* 2006;20:380–390. https://doi.org/10.1016/j.jvoice.2005.04.007.
8. Yiu EML, Yuen YM, Whitehill T, et al. Reliability and applicability of aerodynamic measures in dysphonia assessment. *Clin Linguist Phon.* 2004;18:463–478. https://doi.org/10.1080/02699200410001703592.
9. Hartl DM, Hans S, Vaissière J, et al. Objective voice quality analysis before and after onset of unilateral vocal fold paralysis. *J Voice.* 2001;15:351–361. https://doi.org/10.1016/S0892-1997(01)00037-6.
10. Dastolfo C, Gartner-Schmidt J, Yu L, et al. Aerodynamic outcomes of four common voice disorders: moving toward disorder-specific assessment. *J Voice.* 2016;30:301–307. https://doi.org/10.1016/j.jvoice.2015.03.017.
11. Gillespie AI, Gartner-Schmidt J, Rubinstein EN, et al. Aerodynamic profiles of women with muscle tension dysphonia/aphonia. *J Speech Lang Hear Res.* 2013;56:481–488. https://doi.org/10.1044/1092-4388(2012/11-0217).
12. Lewandowski A, Gillespie AI. The relationship between voice and breathing in the assessment and treatment of voice disorders. *Perspect ASHA Spec Interest Groups.* 2016;1:94–104. https://doi.org/10.1044/persp1.sig3.94.
13. Liang FY, Yang JS, Mei XS, et al. The vocal aerodynamic change in female patients with muscular tension dysphonia after voice training. *YMVJ.* 2014;28:393.e7–393.e10. https://doi.org/10.1016/j.jvoice.2013.11.010.
14. Shin JE, Nam SY, Yoo SJ, et al. Analysis of voice and quantitative measurement of glottal gap after thyroplasty type I in the treatment of unilateral vocal paralysis. *J Voice.* 2002;16:136–142. https://doi.org/10.1016/S0892-1997(02)00083-8.
15. Tong JY, Sataloff RT. Respiratory function and voice: the role for airflow measures. *J Voice.* 2022;36:542–553. https://doi.org/10.1016/J.JVOICE.2020.07.019.
16. ASHA. Scope of practice in speech-language pathology; 2016. Available at: https://www.asha.org/policy/sp2016–00343/#Domains_of_Professional_Practice. Accessed August 4, 2021.
17. Behrman A. Common practices of voice therapists in the evaluation of patients. *J Voice.* 2005;19:454–469. https://doi.org/10.1016/j.jvoice.2004.08.004.
18. Hartley NA, Braden M, Thibeault SL. Practice patterns of speech-language pathologists in pediatric vocal health. *Am J Speech Lang Pathol.* 2017;26:281–300. https://doi.org/10.1044/2016_AJSLP-15-0057.
19. Cohen W, Wynne DMG, Kubba H, et al. Development of a minimum protocol for assessment in the paediatric voice clinic. Part 1: evaluating vocal function. *Logop Phoniatr Vocol.* 2012;37:33–38. https://doi.org/10.3109/14015439.2011.638670.
20. Lee JG, Bang SH, Jeon JM, et al. Study on assessment and treatment patterns of speech-language pathologists in pediatric vocal problem through multicenter survey. *J Korean Soc Laryngol Phoniatr Logop.* 2019;30:39–47.https://jkslp.org/journal/view.php?number=2111.
21. MacBean N, Ford H, Madill C, et al. Common practices of Australian SLPs in the management of paediatric vocal health. Paper presented at: The Voice Foundation, Philadelphia, PA. Published online 2014.

22. McAlister S, Yanushevskaya I. Voice assessment practices of speech and language therapists in Ireland. *Clin Linguist Phon.* 2020;34:29–53. https://doi.org/10.1080/02699206.2019.1610798.
23. Hirano M, Koike Y, Von Leden H. Maximum phonation time and air usage during phonation. Clinical study. *Folia Phoniatr (Basel).* 1968;20:185–201.Accessed September 30, 2014. http://www.ncbi.nlm.nih.gov/pubmed/5671971.
24. Iwata S, von Leden H. Phonation quotient in patients with laryngeal diseases. *Folia Phoniatr Logop.* 1970;22:117–128. https://doi.org/10.1159/000263375.
25. Yoshioka H, Sawashima M, Hirose H, et al. Clinical evaluation of air usage during phonation. *Jpn J Logop Phoniatr.* 1977;18:87–93. https://doi.org/10.5112/jjlp.18.87.
26. Rau D, Beckett RL. Aerodynamic assessment of vocal function using hand-held spirometers. *J Speech Hear Disord.* 1984;49:183–188. https://doi.org/10.1044/jshd.4902.183.
27. Yanagihara N, Koike Y, von Leden H. Phonation and respiration: function study in normal subjects. *Folia Phoniat Logop.* 1966;18:323–340. https://doi.org/10.1159/000263110.
28. Joshi A, Watts CR. Measurement reliability of phonation quotient derived from three aerodynamic instruments. *J Voice.* 2016;30:773.e13–773.e19. https://doi.org/10.1016/j.jvoice.2015.11.015.
29. Joshi A, Watts CR. A comparison of indirect and direct methods for estimating transglottal airflow rate. *J Voice.* 2018;32:655–659. https://doi.org/10.1016/j.jvoice.2017.09.020.
30. Joshi A, Watts CR, Hathway J. Phonation quotient using three aerodynamic instruments in the disordered voice. *J Voice.* 2020;34:20–24. https://doi.org/10.1016/j.jvoice.2018.08.002.
31. Aghajanzadeh M, Darouie A, Dabirmoghaddam P, et al. The relationship between the aerodynamic parameters of voice and perceptual evaluation in the Iranian population with or without voice disorders. *J Voice.* 2017;31:250.e9–250.e15. https://doi.org/10.1016/j.jvoice.2016.07.014.
32. Joshi A, Watts CR. Phonation quotient in women: a measure of vocal efficiency using three aerodynamic instruments. *J Voice.* 2017;31:161–167. https://doi.org/10.1016/j.jvoice.2016.06.007.
33. Morsomme D, Jamart J, Boucquey D, et al. Presbyphonia: voice differences between the sexes in the elderly. Comparison by maximum phonation time, phonation quotient and spectral analysis. *Logop Phoniatr Vocol.* 2009;22:9–14. https://doi.org/10.3109/14015439709075310.
34. Awan S. The aging female voice: acoustic and respiratory data. *Clin Linguist Phon.* 2006;20:171–180. https://doi.org/10.1080/02699200400026918.
35. Vonnegut B. A vortex whistle. *J Acoust Soc Am.* 1954;26:18. https://doi.org/10.1121/1.1907282.
36. Kristiansen U, Amielh M. Measurements on a little known sound source-the Vortex Whistle. In: Viggen E, Kristiansen U, eds. *Proceedings of the 39th Scandinavian Symposium on Physical Acoustics.* Geilo, Norway: Norwegian Physical Society; 2016:1–13.
37. Sato H, Ohara M, Watanabe K, et al. Application of the vortex whistle to the spirometer. *Trans Soc Instrum Control Eng.* 1999;35:840–845. https://doi.org/10.9746/sicetr1965.35.840.
38. Awan SN, Awan JA. Use of a vortex whistle for measures of respiratory capacity. *J Voice.* 2022;36:630–636. https://doi.org/10.1016/j.jvoice.2020.07.038.
39. Awan SN, Awan JA. Comparison of methods of eliciting vital capacity: forced versus slow vital capacity. *J Voice.* 2025. https://doi.org/10.1016/J.JVOICE.2024.12.039. S0892-1997(24)00474-0, Epub ahead of print.
40. R Core Team. R: a language and environment for statistical computing. Published online 2020. Available at: https://www.r-project.org/. Accessed January 2, 2025.
41. Akoglu H. User's guide to correlation coefficients. *Turk J Emerg Med.* 2018;18:91–93. https://doi.org/10.1016/j.tjem.2018.08.001.
42. Steiger J. Tests for comparing elements of a correlation matrix. *Psychol Bull.* 1980;87:245–251.
43. Fisher R. *Statistical Methods for Research Workers*. 11th ed. London, UK: Oliver and Boyd; 1950.
44. Cocor - comparing correlations. Available at: http://comparingcorrelations.org/. Accessed April 30, 2025.
45. Diedenhofen B, Musch J. Cocor: a comprehensive solution for the statistical comparison of correlations. *PLoS One.* 2015;10:1–14. https://doi.org/10.1371/JOURNAL.PONE.0121945.
46. JASP Team. JASP (Version 0.95.1) [Computer software] 2025. Available at: https://jasp-stats.org/.
47. Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Hillsdale NJ: L. Erlbaum Associates; 1988.
48. Hilbe J. *Negative Binomial Regression*. 2nd ed. Cambridge University Press; 2011.Accessed April 21, 2025 www.cambridge.org.
49. Yanagihara N, Koike Y, von Leden H. Phonation and respiration. *Folia Phoniatr Logop.* 1966;18:323–340. https://doi.org/10.1159/000263110.
50. Higgins M, Netsell R, Schulte L. Aerodynamic and electroglottographic measures of normal voice production: intrasubject variability within and across sessions. *J Speech Hear Res.* 1994;37:38–45.Accessed February 19, 2013 http://www.ncbi.nlm.nih.gov/pubmed/8170128.
51. Awan SN, Novaleski CK, Yingling JR. Test-retest reliability for aerodynamic measures of voice. *J Voice.* 2013;27:674–684. https://doi.org/10.1016/j.jvoice.2013.07.002.
52. Holmberg EB, Hillman RE, Perkell JS. Glottal airflow and transglottal air pressure measurements for male and female speakers in soft, normal, and loud voice. *J Acoust Soc Am.* 1988;84:511–529. https://doi.org/10.1121/1.396829.
53. Li S, Scherer RC, Wan MX, et al. Intraglottal pressure: a comparison between male and female larynxes. *J Voice.* 2020;34:813–822. https://doi.org/10.1016/j.jvoice.2019.06.005.
54. American Speech-Language Association. Medicare Coding Rules for SLP Services copy; 2019. Available at: https://www.asha.org/practice/reimbursement/medicare/SLP_coding_rules/. Accessed May 3, 2025.