

## ORIGINAL RESEARCH

# Reliability of peak expiratory flow percentage compared to endoscopic grading in subglottic stenosis

Sungjin A. Song MD<sup>1,2</sup>  | Alena Santeerapharp MD<sup>3</sup> | Kanittha Choksawad MD<sup>4</sup> | Ramon A. Franco Jr MD<sup>1,2</sup> 

<sup>1</sup>Department of Otolaryngology, Massachusetts Eye and Ear, Boston, Massachusetts

<sup>2</sup>Department of Otolaryngology, Harvard Medical School, Boston, Massachusetts

<sup>3</sup>Department of Otorhinolaryngology, Faculty of Medicine, Srinakharinwirot University, Bangkok, Thailand

<sup>4</sup>Department of Otolaryngology, Panyanantapikku Chonprathan Medical Center Srinakharinwirot University, Bangkok, Thailand

## Correspondence

Ramon A. Franco Jr, MD, Department of Otolaryngology, Massachusetts Eye and Ear, 243 Charles Street, Boston, MA 02114. Email: ramon\_franco@meei.harvard.edu

## Funding information

Massachusetts Eye and Ear

## Abstract

**Objective:** To determine the reliability of pulmonary function testing compared to endoscopic grading in the assessment of subglottic stenosis.

**Methods:** Consecutively treated patients with subglottic stenosis at a tertiary care specialty hospital from 2009 to 2019 were identified. Two fellowship-trained laryngologists and two otolaryngologists blinded to clinical history reviewed laryngo tracheoscopic examinations and assessed the degree of stenosis using the Cotton-Myer grading system (% stenosis). Nine full flow-volume loops were performed at the time of each exam.

**Results:** The endoscopic images of 45 subjects were graded for degree of stenosis and the spirometry data were analyzed. The kappa values for Cotton-Myer grade overall was 0.37, grade I was  $-0.103$ , grade II was 0.052, and grade III was 0.045. The overall intraclass correlation of the physician grading of estimated percent obstruction (% stenosis) was 0.712 ( $P < .01$ ) whereas the overall intraclass correlation for PEF% was 0.96 ( $P < .01$ ). Within each Cotton-Myer grade, the intraclass correlation for % stenosis was 0.45 ( $P = .02$ ) for grade I, 0.06 ( $P = .30$ ) for grade II, and 0.16 ( $P = .03$ ) for grade III. The intraclass correlation for PEF% for grade I was 0.97 ( $P < .01$ ), grade II was 0.92 ( $P < .01$ ), and grade III was 0.96 ( $P < .01$ ).

**Conclusion:** Cotton-Myer grading and estimating percent obstruction (% stenosis) for adult subglottic stenosis showed poor reliability as an assessment tool compared to the excellent intraclass correlation seen with pulmonary function tests within each Cotton-Myer grade subgroup. We recommend pulmonary function testing, specifically PEF% because it is a normalized value, for the assessment and management of subglottic stenosis.

**Level of evidence:** 4.

Institution where the work was primarily performed: Massachusetts Eye and Ear, Massachusetts, USA.

This manuscript has not been presented at a meeting. All authors met the criteria for authorship established by the International Committee of Medical Journal Editors, specifically: **Sungjin A. Song** and **Ramon A. Franco** were responsible for substantial contribution to the conception, design, and drafting the work, revising the work and reviewing the manuscript. **Alena Santeerapharp** and **Kanittha Choksawad** were responsible for substantial contribution to the data collection and analysis, drafting the work, revising the work and reviewing the manuscript. Additionally, all authors provided final approval of the version to be published and agreed to be accountable for all aspects of the work in ensuring including the accuracy and/or integrity of the work.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2020 The Authors. *Laryngoscope Investigative Otolaryngology* published by Wiley Periodicals LLC. on behalf of The Triological Society.

## KEYWORDS

Cotton-Myer grading, laryngotracheal stenosis, pulmonary function testing, subglottic stenosis

## 1 | INTRODUCTION

Subglottic stenosis is a narrowing of the upper airway that can result in severe dyspnea. Usually, physicians will make management decisions for subglottic stenosis based on the clinical history and the degree of stenosis using subjective visual endoscopic assessment. Patients often present for treatment with advanced stenosis and severe dyspnea. Open techniques such as cricotracheal resection and laryngotracheal reconstruction are associated with high perioperative morbidity and poor long-term voice outcomes. With the advances in lasers and airway balloons, less invasive endoscopic approaches (balloon dilation, endoscopic resection, serial intralesional steroid injections [SILSI], etc.) have become the primary treatment for subglottic stenosis reserving open surgery for recalcitrant cases.<sup>1,2</sup>

The Cotton-Myer grading system is one of the most widely used tools in evaluating the severity of subglottic stenosis.<sup>3</sup> The four grades are categorized as: 0% to 50% for grade I, 51% to 70% for grade II, 71% to 99% for grade III, and complete luminal obstruction for grade IV.<sup>4</sup> Current studies use the subjective Cotton-Myer grade or estimated percent obstruction (% stenosis) as the exclusive objective outcome measure in the management of adult subglottic stenosis.<sup>5-7</sup> These studies assume that endoscopic grading is sufficient in following patients with subglottic stenosis. However, Murgu et al<sup>8</sup> found that experienced bronchoscopists misclassified the degree of laryngotracheal stenosis 53% of the time compared to morphometric bronchoscopic measurements. The majority of misclassifications (684 of 755 [91%]) were under-classifications. The ability to simply assess the degree of stenosis based on endoscopic assessment may not be as reliable a method as previously assumed.

In subglottic stenosis, pulmonary function testing (PFT) can reveal functional changes related to disease progression and treatment response. In 1973, Miller and Hyatt<sup>9</sup> reported that plateauing of the highest portion of the forced expiratory loop is the first to be distorted on flow volume loop with simulated upper airway obstruction. Nouraei et al<sup>10</sup> was one of the first to examine the utility of spirometry indices (total peak flow and ratio of area under the flow-volume loop to forced vital capacity) in providing an objective, quantitative outcome for treatment-related changes in subglottic stenosis patients. Recent studies have used various spirometry measurements, namely peak inspiratory flow (PIF),<sup>11-13</sup> peak expiratory flow (PEF),<sup>3,13,14</sup> peak expiratory flow percentage (PEF%),<sup>15,16</sup> and total peak flow (TPF)<sup>10</sup> in subglottic stenosis patients to assess disease progression and treatment response. By detecting early aerodynamic decline with PFT before severe clinical symptoms become apparent, laryngologists can start treating patients before the stenosis becomes critical and life-threatening.

In order to better understand the reliability of endoscopic grading compared to PFT in the management of subglottic stenosis, we performed a blinded, prospective grading of nasolaryngoscopic and tracheoscopic examinations of subglottic stenosis and compared the Cotton-Myer grade and estimated degree of stenosis with PFT values. With two different methods to determine stenosis, the central question we have is which system is more reliable to assess the stenosis—the Cotton-Myer grading system (% stenosis) or PFTs (specifically PEF%).

## 2 | MATERIALS AND METHODS

The Partners institutional review board approved this study protocol (2019P001694). A database containing consecutively entered subglottic stenosis patients between January 1, 2009 and January 1, 2020 at Massachusetts Eye and Ear was maintained by the senior laryngologist (R.A.F.). Inclusion criteria were adults ( $\geq 18$  years old) with subglottic stenosis on flexible distal-chip video nasolaryngoscopy and tracheoscopy with three complete forced spirometry tests performed with the handheld Carefusion MicroLoop spirometer (Vyair, Mettawa, IL) at the time of endoscopic assessment. Exclusion criteria were children ( $< 18$  years old), presence of a tracheostomy tube, vocal fold paralysis, no video nasolaryngoscopy/tracheoscopy exam, and lack of complete pulmonary function testing data.

We created our patient reference library, who we would evaluate using both the Cotton-Myer grading system (% stenosis) and PFTs, using our inclusion and exclusion criteria. Data extracted included age, sex, diagnosis, nasolaryngoscopic and tracheoscopic images, and PFT data, specifically peak expiratory flow (PEF), peak inspiratory flow (PIF), peak expiratory flow percentage (PEF%), and total peak flow (TPF). TPF is the sum of PEF and PIF (TPF=PEF + PIF). PFTs were performed using a handheld spirometer in accordance with typical protocols for forced spirometry. PFTs were considered complete if there were 3 separate runs of 3 individual breathing trials (9 separate flow-volume loops). The best PEF, PIF, PEF% values from each run were recorded giving each subject 3 separate values of each parameter to evaluate.

Two fellowship-trained laryngologists (R.A.F. and S.A.S.) and two otolaryngologists (A.S. and K.C.) who were blinded to the clinical history and pulmonary function testing data reviewed flexible distal-chip nasolaryngoscopic and tracheoscopic examinations performed by the senior laryngologist (R.A.F.). They individually determined the percent stenosis (% stenosis) for each patient ranging from no stenosis of 0% to complete obstruction of 100%. The intraclass correlation for % stenosis was then compared with PFT indices. For subgroup analysis, the mean % stenosis was used to categorize the patient into the proper Cotton-Myer grade.

## 2.1 | Statistical analysis

All data were analyzed using SPSS Statistics (version 25.0; IBM). The categorical variables were reported as numbers (n) and proportions (%). The Fleiss' kappa was calculated and reported to assess the interrater reliability for categorical ratings (Cotton-Myer Grading). Kappa values were interpreted based on the following guidelines: <0 poor agreement, .0 to .20 slight agreement, .21 to .40 fair agreement, .41 to .60 moderate agreement, .61 to .80 substantial agreement, and .81 to 1.00 almost perfect agreement.<sup>17</sup> The more commonly used and accepted two-way mixed effect model was used for calculating the intraclass correlation. The intraclass correlation was reported for the correlation between each physician grading of % stenosis, and three consecutive pulmonary function tests for each patients. Intraclass correlation was evaluated using the following guidelines: values less than 0.5 are indicative of poor reliability, values between 0.5 and 0.75 indicate moderate reliability, values between 0.75 and 0.9 indicate good reliability, and values greater than 0.90 indicate excellent reliability.<sup>18</sup> P-values  $\leq .05$  were determined to be statistically significant.

## 3 | RESULTS

We identified 45 patients with the diagnosis of subglottic stenosis confirmed by flexible nasolaryngoscopy and tracheoscopy who met inclusion and exclusion criteria. A total of 45 endoscopic images of subglottic stenosis were graded for degree of stenosis (Table 1). The associated PFTs at the time of each exam were collected and analyzed. The interrater reliability of Cotton-Myer grading showed that overall there was fair agreement, but within each Cotton-Myer grade there was poor or slight agreement (Table 2). The overall intraclass

**TABLE 1** Patient characteristics of included subglottic stenosis patients (n = 45)

Characteristic	n (%)
<b>Gender</b>	
Female	44 (97.8)
Male	1 (0.2)
<b>Age (years)</b>	
Mean ( $\pm$ SD)	50.33 ( $\pm$ 13.96)
Range	18-73
<b>Etiology</b>	
Idiopathic	44 (97.8)
GPA	1 (0.2)
Trauma	0
Other	0
<b>Smoking history</b>	
Never smoker	40 (88.9)
Former smoker	5 (11.1)
Current smoker	0

correlation of the physician grading of % stenosis was 0.712 ( $P < 0.01$ ) and 0.96 ( $P < 0.01$ ) for PEF% (Table 3). Within each Cotton-Myer grade, the intraclass correlation for % stenosis was 0.45 ( $P = .02$ ) for grade I, 0.06 ( $P = .30$ ) for grade II, and 0.16 ( $P = .03$ ) for grade III. The intraclass correlation for PEF% for grade I was 0.97 ( $P < .01$ ), grade II was 0.92 ( $P < .01$ ), and grade III was 0.96 ( $P < .01$ ). The intraclass correlation for total peak flow (TPF) for grade I was 0.98 ( $P < .01$ ), grade II was 0.99 ( $P < .01$ ), and grade III was 0.98 ( $P < .01$ ). Figures 1 through 3 demonstrate the increased variability of the Cotton-Myer grading system (% stenosis) compared to the PEF% in Cotton-Myer grades I through III subglottic stenosis.

## 4 | DISCUSSION

Although pulmonary function testing (PFT) requires maximal respiratory effort and is therefore patient effort dependent, the results from our study show that there is a high level of consistency despite these limitations. Our study found excellent overall intraclass correlation with all PFT measures: PEF was 0.97 ( $P < .01$ ), PIF was 0.96 ( $P < .01$ ), PEF% was 0.96 ( $P < .01$ ) and TPF was 0.99 ( $P < .01$ ). Moreover, the excellent intraclass correlation for PFT values was seen within each Cotton-Myer grade group as well. For example, the intraclass correlation for the Cotton-Myer grade I stenosis group was 0.97 for PEF ( $P < .01$ ), 0.94 for PIF ( $P < .01$ ), 0.97 for PEF% ( $P < .01$ ), and 0.98 for TPF ( $P < .01$ ). PFT improves the resolution and precision with which patients and laryngologists can track disease progression and treatment response. For example, if a patient is at 30% stenosis and progresses to 50%, the Cotton-Myer grading system is unable to resolve this change (grade I stenosis  $\rightarrow$  grade I stenosis), but the decline in cross-sectional airway can be detected with PFT. PFTs allows the clinician to appreciate changes in stenosis before the onset of dyspnea which may influence timing of follow-up and treatment. The increased precision in subglottic stenosis assessment can lead to earlier detection of subclinical narrowing.

Peak expiratory flow percentage (PEF%) can be used to compare between patients and management strategies. The intraclass correlation for all the PFT values (PEF, PIF, PEF%, and TPF) were similar and categorized as excellent. Notably, there was excellent intraclass correlation for PEF% for all Cotton-Myer grade groups: grade I was 0.97 ( $P < .01$ ), grade II was 0.92 ( $P < .01$ ), and grade III was 0.96 ( $P < .01$ ). We record PEF, PIF, and PEF% for all of our subglottic stenosis

**TABLE 2** Interrater reliability of Cotton-Myer grading of 45 subglottic stenosis endoscopic images

	Fleiss' kappa	95% CI	P-value
Overall	0.370	0.283-0.456	<.01
Cotton-Myer I	-0.103	-0.386-0.179	0.474
Cotton-Myer II	0.052	-0.140-0.244	0.597
Cotton-Myer III	0.045	-1.01-0.191	0.547

Abbreviation: CI, confidence interval.

**TABLE 3** Intraclass correlation of 45 endoscopic images of subglottic stenosis and corresponding pulmonary function tests (PFTs). Estimated percent obstruction (% stenosis) was the same as the Cotton-Myer grading system. Here we see the variability of using the Cotton-Myer grading system between the different physician graders vs PFTs. % stenosis has a lower intraclass correlation compared to any of the PFT parameters (PEF, PIF, PEF%, TPF), which have excellent reliability both overall and within each Cotton-Myer grade group

	Correlation <sup>a</sup>	95% CI	P-value
<b>Overall</b>			
% stenosis	0.712	0.601-0.808	<.01
PEF	0.966	0.945-0.980	<.01
PIF	0.958	0.933-0.975	<.01
PEF%	0.956	0.930-0.974	<.01
TPF	0.985	0.975-0.991	<.01
<b>Cotton-Myer I</b>			
% stenosis	0.450	0.123-0.782	.02
PEF	0.970	0.904-0.993	<.01
PIF	0.944	0.827-0.988	<.01
PEF%	0.970	0.903-0.993	<.01
TPF	0.984	0.948-0.997	<.01
<b>Cotton-Myer II</b>			
% stenosis	0.061	-0.150-0.468	.303
PEF	0.968	0.905-0.992	<.01
PIF	0.949	0.851-0.987	<.01
PEF%	0.915	0.763-0.978	<.01
TPF	0.987	0.961-0.997	<.01
<b>Cotton-Myer III</b>			
% stenosis	0.159	-0.001-0.375	.026
PEF	0.963	0.931-0.981	<.01
PIF	0.963	0.932-0.981	<.01
PEF%	0.962	0.930-0.981	<.01
TPF	0.984	0.971-0.992	<.01

Abbreviations: PEF, peak expiratory flow; PEF%, peak expiratory flow percentage; PIF, peak inspiratory flow, TPF; total peak flow.

<sup>a</sup>Intraclass correlation: Values less than 0.5—poor; between 0.5 and 0.75—moderate; between 0.75 and 0.9—good; greater than 0.90—excellent reliability.

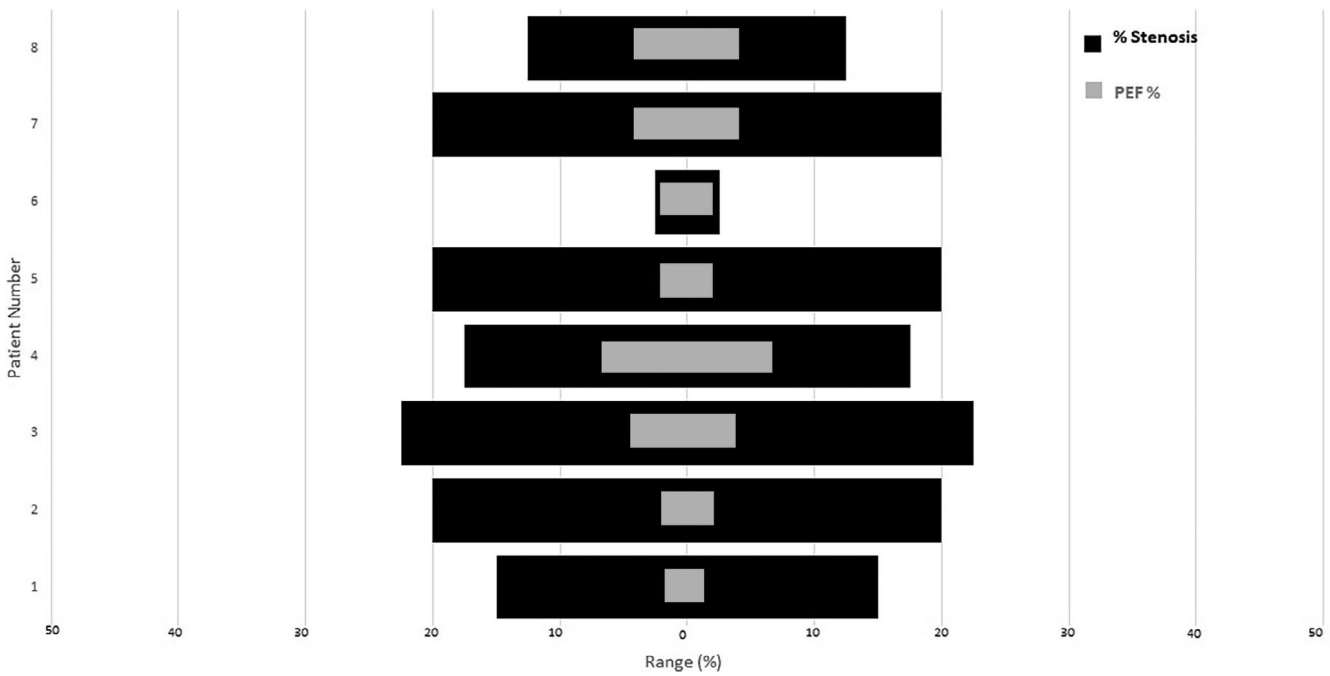
patients at each visit and recommend using PEF%, because it is a normalized value that can be used to compare between patients. Absolute values such as PEF, PIF, and TPF are not optimal to compare patients because the normal values vary for each patient based on sex, age, height, race, and smoking status.<sup>19</sup> However, in a relatively homogeneous group such as idiopathic subglottic stenosis, PEF alone has been shown to work well to track individual patients.<sup>14</sup> PEF% is a very highly reliable PFT value that allows surgeons to compare treatment outcomes between different patients, which leads to improving evidence-based treatment for this rare disease.

Our study found that surgeons are able to distinguish between obvious differences in stenosis, but are less likely to appreciate subtle

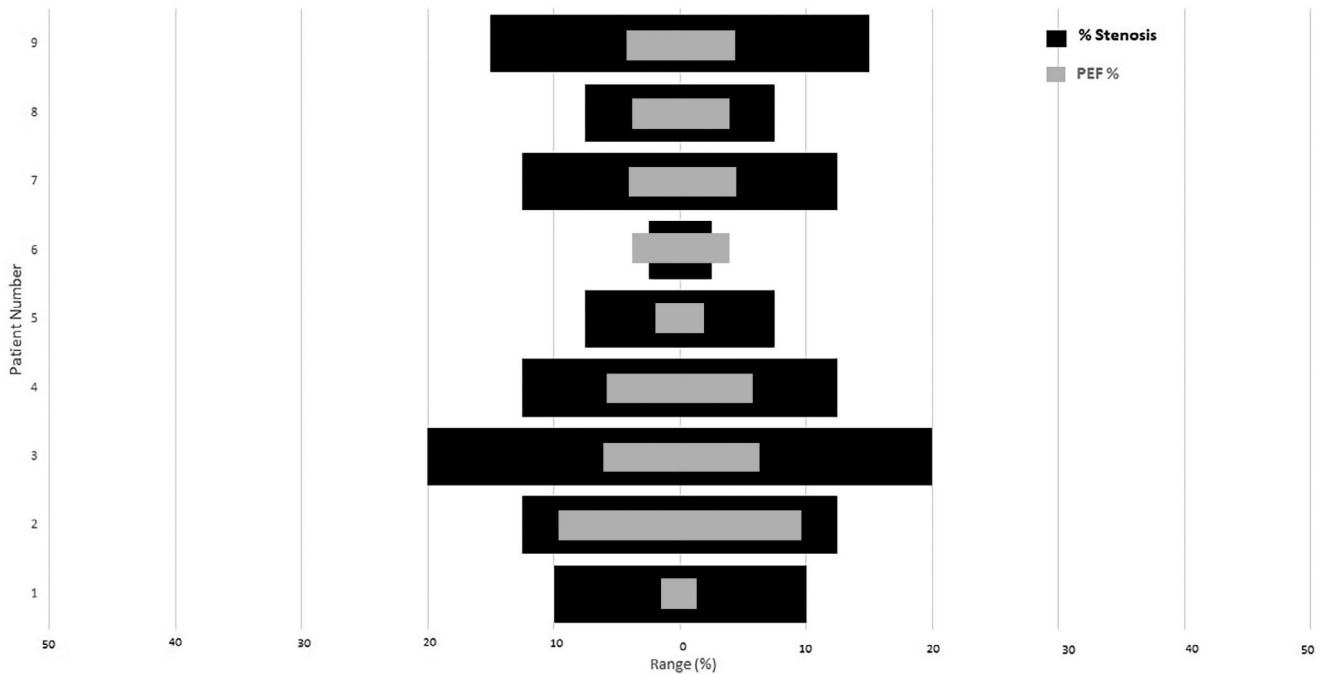
changes. The interrater reliability of Cotton-Myer grading showed that overall there was fair agreement, but within each Cotton-Myer grade there was poor or slight agreement (Table 2). The Cotton-Myer grading of subglottic stenosis had moderate intraclass correlation as a whole but there was very low inter-observer concordance within individual grades. Although the intraclass correlation overall was 0.71 ( $P < .01$ ), the intraclass correlation was poor in each individual Cotton-Myer grade group: grade I was 0.45 ( $P = .02$ ), grade II was 0.06 ( $P = .30$ ), and grade III was 0.16 ( $P = .03$ ). In 1984, Cotton<sup>20</sup> originally proposed a grading scale for pediatric subglottic stenosis based on the perceived percentage of obstruction and the anatomic location. Myer et al<sup>4</sup> then recognized that the subjective interpretation of airway size was imprecise and often inaccurate. Consequently, the Cotton-Myer grading scale used endotracheal tubes to size the airway to determine the degree of stenosis instead of subjective endoscopic assessment.<sup>21</sup> This grading scale and the concept of visually estimating the degree of stenosis was adopted out of convenience and necessity in adults for lack of a widely adopted alternative solution. As we continue to progress to advocating for early detection and treatment of subglottic stenosis in the office setting, we need more objective and reliable ways to determine airway compromise in the eccentric, irregular and corkscrew-shaped stenoses often seen in adults with idiopathic subglottic stenosis. We also need a system that is reliable, highly reproducible and is sensitive to change not perceptible to endoscopic visualization.

Several factors that may limit precise endoscopic assessment of subglottic stenosis include the inability to appreciate the intraluminal cricoid cartilage border (ie, where is the cartilage?), variable distance between the tip of the scope and the stenosis, and image distortion. In subglottic stenosis, inflammation and fibrosis can obscure the endoscopic visualization of the cricoid cartilage. Consequently, it can be difficult to appreciate the border of the cricoid cartilage to accurately and precisely estimate the degree of stenosis. Masters et al<sup>22</sup> reported that length, perimeter and area measurements can be altered based on a target's location within the field of view and distance from the tip of endoscope. For example, peripheral objects appear smaller than central ones and the distortion increases when the lens-to-object distance is reduced.<sup>23</sup> Often to completely view the subglottic stenosis, the tip of the endoscope needs to pass the true vocal folds reducing the distance from the lens to the stenosis. The distortion increases exponentially when within 1 cm from the stenosis making appreciation of the size of stenosis very difficult even when the distance from the stenosis is known.<sup>22</sup> Nouraei et al<sup>23</sup> developed an objective sizing approach that adjusted for tilt, radial, and distance distortions by using a circular calibration probe under general anesthesia during intraoperative direct laryngoscopy and tracheoscopy. However, placement of a calibrator probe at the level of the stenosis in the clinic setting during flexible nasolaryngoscopy and tracheoscopy would be challenging. Subjective grading of subglottic stenosis is also limited by the clarity and resolution of the examination and physician's experience.<sup>21</sup>

Imprecise grading, such as underestimation, can lead to delayed care for patients who may progress to critical levels of stenosis while



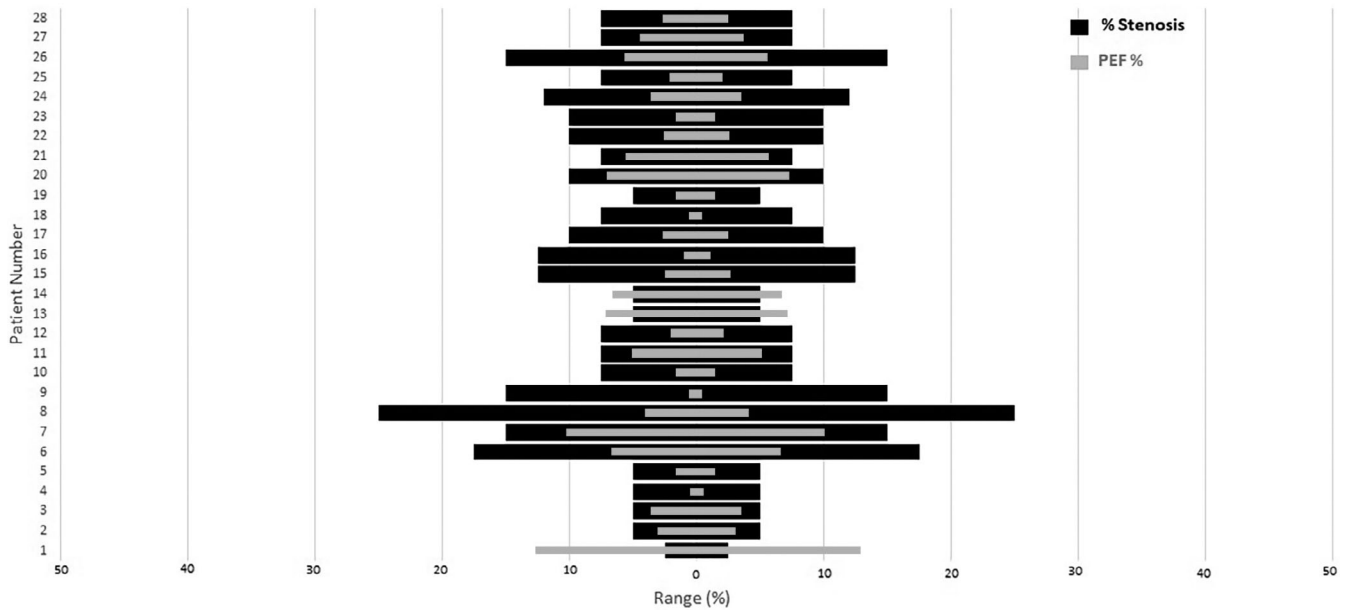
**FIGURE 1** The variability of the estimated percent obstruction (% stenosis) and peak expiratory flow percentage (PEF%) of Cotton-Myer grade I patients. The superior reliability of PEF% is seen in that the spread of the PEF% for each patient is much smaller and tighter (gray) compared to the much larger range (wider bars) for % stenosis (black)



**FIGURE 2** The variability of the estimated percent obstruction (% stenosis) and peak expiratory flow percentage (PEF%) of Cotton-Myer grade II patients. The superior reliability of PEF% is seen in that the spread of the PEF% for all but one patient is much smaller and tighter (gray) compared to the much larger range (wider bars) for % stenosis (black)

overestimations may beget unnecessary and over-aggressive surgeries. Airflow dynamic studies show that a 50% reduction in airway lumen cross-sectional area results in a pressure drop of the same

magnitude as the normal glottic opening, and hence Cotton-Myer grade I (<50% stenosis) is unlikely to lead to significant clinical symptoms.<sup>24</sup> Since resistance is inversely proportional to the radius to the



**FIGURE 3** The variability of the estimated percent obstruction (% stenosis) and peak expiratory flow percentage (PEF%) of Cotton-Myer grade III patients. The superior reliability of PEF% is seen in that the spread of the PEF% for all but 3 patients is much smaller and tighter (gray) compared to the much larger range (wider bars) for % stenosis (black)

fourth power, small changes within a Cotton-Myer grade can result in major changes in airway resistance.<sup>23,25</sup> Moreover, the resistance increases exponentially for the same reduction of cross sectional area when the stenosis is narrower. Because patients may not experience significant symptoms until their stenosis has reached a critical level, they may present after their stenosis has progressed to a higher grade. Precise assessment is crucial to be able to monitor changes earlier in disease progression.

The PEF% (pulmonary function test) is a better measuring tool than endoscopic grading in tracking and comparing treatment effectiveness in subglottic stenosis patients. Although there are four grades in the Cotton-Myer grading scale, effectively only grades I through III are used in patients without a tracheotomy. Thus, after excluding grade IV (complete luminal obstruction), there are only three utilized Cotton-Myer grades. Thus, the Cotton-Myer grading scale when compared to PFT indices, such as the PEF%, is analogous to measuring an object using a three notched ruler vs a 100 notched ruler. The ability to appreciate the differences between two data points with precision is naturally easier with a measuring instrument that has smaller measurement increments. In addition, at-home peak flow meters can be used to allow patients to monitor changes between clinic visits. Close monitoring between clinic visits then increases the ability for the patient to detect and respond to changes in their airway caliber before breathing problems become severe. PFTs are more sensitive to changes and produce a more reliable assessment compared to subjective endoscopic grading.

The limitations of this study include the small number of laryngologists grading the subglottic stenosis. Although, the physicians grading the stenoses were blinded when viewing the

examinations, the senior author (R.A.F.) performed the flexible nasolaryngoscopy and tracheoscopy at the time the examinations were recorded. There was no control group without stenosis or repeated examinations of the same patient to assess for intra-observer reliability for Cotton-Myer grading. The objective of the study was to determine the reliability of the Cotton-Myer grading scale and pulmonary function testing as measurement tools. Determining the intra-observer reliability and generalizability of these findings will be studied in the future.

## 5 | CONCLUSION

This preliminary study provides data showing superior reliability of pulmonary function testing compared to the low interobserver concordance with Cotton-Myer grading and estimated percent obstruction (% stenosis) in adult subglottic stenosis. All PFT indices analyzed (PEF, PIF, PEF%, and TPF) had excellent intraclass correlation. However, we prefer the use of PEF% because it is a normalized value that allows us to compare the results between patients.

## ACKNOWLEDGMENTS

The authors acknowledge Hui Zheng, PhD for statistical support and guidance provided during the execution of this project. This work was conducted with support from Harvard Catalyst Biostatistical Consulting Program.

## CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

**DISCLAIMER**

The views expressed in this manuscript are those of the authors and do not reflect the official policy or position of the Department of the Army, Department of Defense or the U.S. Government.

**ORCID**

Sungjin A. Song  <https://orcid.org/0000-0002-4335-1236>

Ramon A. Franco Jr  <https://orcid.org/0000-0002-4549-6017>

**REFERENCES**

- Gelbard A, Anderson C, Berry LD, et al. Comparative treatment outcomes for patients with idiopathic subglottic stenosis. *JAMA Otolaryngol Head Neck Surg*. 2019;31:1-10.
- Nouraei SA, Singh A, Patel A, Ferguson C, Howard DJ, Sandhu GS. Early endoscopic treatment of acute inflammatory airway lesions improves the outcome of postintubation airway stenosis. *Laryngoscope*. 2006;116(8):1417-1421.
- Abdullah A, Alrabiah A, Habib SS, et al. The value of spirometry in subglottic stenosis. *Ear Nose Throat J*. 2019;98(2):98-101.
- Myer CM 3rd, O'Connor DM, Cotton RT. Proposed grading system for subglottic stenosis based on endotracheal tube sizes. *Ann Otol Rhinol Laryngol*. 1994;103(4 Pt 1):319-323.
- Damrose EJ, Manson L, Nekhendzy V, Collins J, Campbell R. Management of subglottic stenosis in pregnancy using advanced apnoeic ventilatory techniques. *J Laryngol Otol*. 2019;133(5):399-403.
- Hoffman MR, Francis DO, Mai JP, Dailey SH. Office-based steroid injections for idiopathic subglottic stenosis: patient-reported outcomes, effect on stenosis, and side effects. *Ann Otol Rhinol Laryngol*. 2019;20:3489419889066.
- Hoffman MR, Coughlin AR, Dailey SH. Serial office-based steroid injections for treatment of idiopathic subglottic stenosis. *Laryngoscope*. 2017;127(11):2475-2481.
- Murgu S, Colt H. Subjective assessment using still bronchoscopic images misclassifies airway narrowing in laryngotracheal stenosis. *Interact Cardiovasc Thorac Surg*. 2013;16(5):655-660.
- Miller RD, Hyatt RE. Evaluation of obstructing lesions of the trachea and larynx by flow-volume loops. *Am Rev Respir Dis*. 1973;108(3):475-481.
- Nouraei SM, Franco RA, Dowdall JR, et al. Physiology-based minimum clinically important difference thresholds in adult laryngotracheal stenosis. *Laryngoscope*. 2014;124(10):2313-2320.
- Yung KC, Chang J, Courey MS. A randomized controlled trial of adjuvant mitomycin-c in endoscopic surgery for laryngotracheal stenosis. *Laryngoscope*. 2019;130(3):706-711.
- Tasche KK, Bayan S, Schularick NM, Wilson J, Hoffman HT. Utility of peak inspiratory flow in managing subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2015;124(6):499-504.
- Kraft SM, Sykes K, Palmer A, Schindler J. Using pulmonary function data to assess outcomes in the endoscopic management of subglottic stenosis. *Ann Otol Rhinol Laryngol*. 2015;124(2):137-142.
- Carpenter DJ, Ferrante S, Bakos SR, Clary MS, Gelbard AH, Daniero JJ. Utility of routine spirometry measures for surveillance of idiopathic subglottic stenosis. *JAMA Otolaryngol Head Neck Surg*. 2019;145(1):21-26.
- Franco RA Jr, Husain I, Reder L, Paddle P. Awake serial intralesional steroid injections without surgery as a novel targeted treatment for idiopathic subglottic stenosis. *Laryngoscope*. 2018;128(3):610-617.
- Naunheim MR, Paddle PM, Husain I, Wangchalabovorn P, Rosario D, Franco RA Jr. Quality-of-life metrics correlate with disease severity in idiopathic subglottic stenosis. *Laryngoscope*. 2018;128:1398-1402.
- Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33(1):159-174.
- Koo TK, Li MY. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J Chiropr Med*. 2016;15(2):155-163.
- Song SA, Sandhu G, Franco RAJ. Should we routinely use pulmonary function testing in the management of subglottic stenosis? *Laryngoscope*. 2020. <https://doi.org/10.1002/lary.28678>.
- Cotton RT. Pediatric laryngotracheal stenosis. *J Pediatr Surg*. 1984;19(6):699-704.
- Murgu S, Colt HG. Morphometric bronchoscopy in adults with central airway obstruction: case illustrations and review of the literature. *Laryngoscope*. 2009;119(7):1318-1324.
- Masters IB, Eastburn MM, Francis PW, et al. Quantification of the magnification and distortion effects of a pediatric flexible video-bronchoscope. *Respir Res*. 2005;6:16.
- Nouraei SA, McPartlin DW, Nouraei SM, et al. Objective sizing of upper airway stenosis: a quantitative endoscopic approach. *Laryngoscope*. 2006;116(1):12-17.
- Brouns M, Verbanck S, Lacor C. Influence of glottic aperture on the tracheal flow. *J Biomech*. 2007;40(1):165-172.
- Bock KR, Silver P, Rom M, Sagy M. Reduction in tracheal lumen due to endotracheal intubation and its calculated clinical significance. *Chest*. 2000;118(2):468-472.

**How to cite this article:** Song SA, Santeerapharp A, Choksawad K, Franco RA Jr. Reliability of peak expiratory flow percentage compared to endoscopic grading in subglottic stenosis. *Laryngoscope Investigative Otolaryngology*. 2020;5:1133-1139. <https://doi.org/10.1002/lio2.492>