Validation of a grading system for lateral nasal wall insufficiency

Gabriel J. Tsao, M.D., Natalia Fijalkowski, and Sam P. Most, M.D.

ABSTRACT

This study was designed to validate a grading scheme for lateral nasal wall insufficiency with interrater and intrarater reliability measures. Representative endoscopic videos depicting varied degrees of lateral nasal wall insufficiency were collated into a 30-clip video (15 clips in duplicate). This was rated by five reviewers for a total of 150 observations. Interrater and intrarater reliability were determined using Fleiss kappa and intraclass correlation coefficient (ICC) statistics, respectively. Good agreement was established between reviewers (interrater reliability), with a Fleiss kappa of 0.7733 ($p < 0.01$). Analysis of intrarater variability with the ICC revealed a very strong agreement (ICC = 0.88; $p < 0.01$). The proposed grading system is shown to have good interrater and intrarater reliability. It provides a reliable instrument for assessing lateral wall insufficiency.

Nasal obstruction is a commonly encountered complaint in otolaryngology practice, and the etiologies of nasal obstruction are many. The classically described surgically treatable causes of nasal airway obstruction are septal deviation, turbinate hypertrophy, and internal and external valve collapse.

The nomenclature of external valve collapse and so-called dynamic internal valve collapse have been discussed previously.[1] Specifically, it has been proposed that all lateral wall insufficiency (LWI) be classified by the zone in which it occurs (Fig. 1). Zone 1 LWI occurs more cephalad and corresponds to dynamic movement of the nasal sidewall at the level of the upper lateral cartilage and scroll region. Zone 2 collapse occurs at the level of the ala and is more akin to classically described external valve collapse.

Bernoulli's principle of fluid dynamics states that as the speed of a moving fluid increases, the pressure within the fluid decreases. Within the nose, air is the fluid medium, and as it increases in speed through the nasal cavity, negative pressure develops, pulling the lateral nasal wall in toward the septum resulting in LWI.

Today, there is no reliability-tested grading scheme for assessing LWI. In the aforementioned article published by the senior author (S.P.M.) in 2008,[1] a grading scheme for zone 1 LWI was proposed. This article aims to validate this grading scheme with interrater and intrarater statistical analysis.

METHODS

Study Design

This study included patients seen at the Stanford University Medical Center, Department of Otolaryngology–Head and Neck Surgery, Facial Plastics Division. Study protocols were approved by the Institutional Review Board at Stanford University. All were patients of the senior author (S.P.M.). Endoscopic videos of the internal nasal valve were taken with a 0° 4-mm Storz-Hopkins telescope through two nasal breath cycles.

The videos were graded by the first author (G.J.T.) according to the schema proposed by their group previously. A description of the grading scheme can be seen in Fig. 2. Fifteen videos were selected—five videos showed each of the three grades of collapse—encompassing the spectrum of lateral nasal wall collapse. Each video was blinded of patient data, duplicated, and then sequence randomized to a total of 30 clips. The duplication was important to allow assessment of intrarater reliability.

A 19-minute video was then compiled of the 30 clips. The raters were allowed to rewind and replay as needed for their grading. Five raters (two fellows and three otolaryngology residents) were asked to grade these clips.

Data Analysis

All statistical analyses were performed with Statistical Analysis Software (SAS) Enterprise Guide Version 4.3 (SAS Institute, Inc., Cary, NC). Interrater agreement was assessed with Fleiss κ to account for the ordinal

From the Division of Facial Plastic and Reconstructive Surgery, Stanford University School of Medicine, Stanford, California
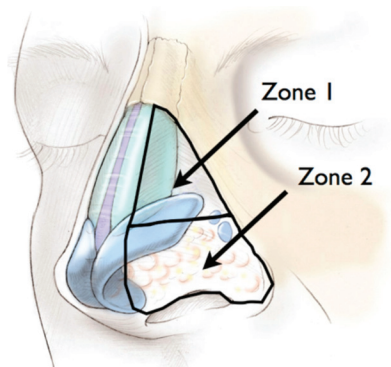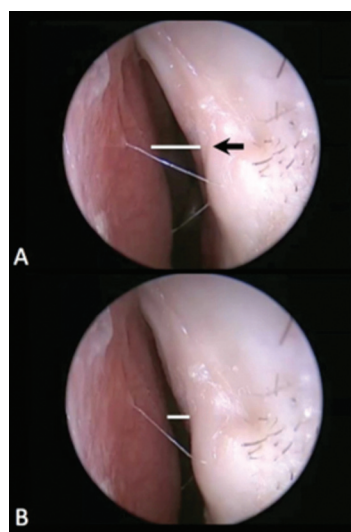
Figure 1. *Zones of lateral nasal wall insufficiency. Zone 1 (upper zone) corresponds to dynamic internal nasal valve collapse. Zone 2 (lower zone) corresponds to classic external valve collapse.*

Figure 2. *Grading system for lateral nasal wall collapse. (1) Identify the junction of the upper and lower cartilages (black arrow) just above the recurvature of the lower lateral cartilage. (2) Imagine a line parallel to the nasal floor across to the septum. (3) Estimate the degree of collapse during inspiration as a percentage compared with repose.*

nature of the grading scale as well as the number of raters. Because the grading scale consisted of more than two categories, the weighted κ takes into account the amount of agreement (otherwise known as partial agreement) on the grading scale. For interpretation of interrater agreement the following standard measures were followed: poor ($\kappa < 0.2$), fair ($\kappa = 0.21–0.4$), moderate ($\kappa = 0.41–0.60$), good ($\kappa = 0.61–0.80$), and very good agreement ($\kappa = 0.81–1.00$).[2] For interrater agreement, only the first observation was included in analysis as determined *a priori* by the investigators. The intraclass correlation coefficient (ICC) was determined to assess the index of reliability for a single rater (averaged for overall statistics across the five raters in this study).

Table 1 **Interrater agreement using the proposed LWI grading system**

Video No.	Grade 1	Grade 2	Grade 3	Fleiss κ (clinical agreement)
1	5	0	0	1.00
2	0	5	0	1.00
3	0	4	1	0.60
4	5	0	0	1.00
5	0	2	3	0.40
6	5	0	0	1.00
7	0	1	4	0.60
8	2	3	0	0.40
9	0	0	5	1.00
10	3	0	2	0.40
11	0	1	4	0.60
12	5	0	0	1.00
13	5	0	0	1.00
14	2	3	0	0.40
15	5	0	0	1.00
Cumulative				$\kappa = 0.77$ (good agreement); $p < 0.01$

Descriptions for degree of clinical agreement are based on the following previously published measures[2]: poor ($\kappa < 0.2$), fair ($\kappa = 0.21–0.4$), moderate ($\kappa = 0.41–0.60$), good ($\kappa = 0.61–0.80$), and very good agreement ($\kappa = 0.81–1.00$). All interrater agreement κ's were calculated using a Fleiss κ-statistic.
LWI = lateral wall insufficiency.

RESULTS

Five observers each reviewed and graded 15 videos twice after completing initial grading scale training for a total of 150 observations. As mentioned previously, the interrater agreement was determined using the first observed video rating for each of the reviewers resulting in 75 independent observations. Table 1 lists the weighted κ for each video as well as the interrater agreement when taking into account all 75 observations. Overall, the Fleiss κ was 0.7733 ($p < 0.01$), indicating good agreement between users. Each reviewer rated the 15 videos twice, given in a random order. Analysis of intrarater variability with the ICC revealed a very strong agreement (ICC = 0.88; $p < 0.01$).

DISCUSSION

This study aimed to present a validated method for grading zone 1 LWI or dynamic internal valve collapse. Nasal obstruction is a common complaint in an otolaryngologist's office, and a standardized assessment of LWI would be of benefit. To date, there is no accepted standard for grading this collapse.

The proposed system is a subjective measurement by design. Objective metrics such as acoustic rhinometry are involved and are costly to the point of being impractical on a clinical level. Visual assessment with this grading scheme is fast, simple, and easily grasped by trainees and patients. Here, we show its validity, with good agreement between different observers (interrater reliability), as well as strong agreement within each observer (intrarater reliability).

This system should be used to assess surgical outcomes when addressing LWI. It can be used in conjunction with the Nasal Obstruction Symptom Evaluation scale—a validated instrument for assessing a patient's symptoms of nasal obstruction.[3] Hopefully, this grading system will help facilitate further study to provide evidence-based guidelines to address LWI.

REFERENCES

1. Most SP. Trends in functional rhinoplasty. Arch Facial Plast Surg 10:410–413, 2008.
2. Brennan P, and Silman A. Statistical methods for assessing observer variability in clinical measures. BMJ 304:1491–1494, 1992.
3. Stewart MG, Smith TL, Weaver EM, et al. Outcomes after nasal septoplasty: Results from the Nasal Obstruction Septoplasty Effectiveness (NOSE) study. Otolaryngol Head Neck Surg 130: 283–290, 2004. □