

# Comparison of different endoscopic scoring systems in patients with chronic rhinosinusitis: reliability, validity, responsiveness and correlation\*

Lu Zhang<sup>1</sup>, Linghao Zhang<sup>2</sup>

<sup>1</sup> Department of Otorhinolaryngology, Ningbo No. 2 Hospital, Ningbo, China

<sup>2</sup> Department of Otorhinolaryngology, The First Affiliated Hospital of Wenzhou Medical University, Wenzhou, China

**Rhinology** 55: 363-368, 2017

<https://doi.org/10.4193/Rhino17.109>

**\*Received for publication:**

May 26, 2017

**Accepted:** June 24, 2017

## Abstract

**Objectives:** Previous studies have proposed various endoscopic scoring systems to assess disease severity of patients with chronic rhinosinusitis (CRS). However, those assessments have not been validated. This study was designed to assess the Modified Lund - Kennedy (MLK) and the discharge, inflammation, and polyps/oedema (DIP) endoscopic scoring systems in patients with CRS.

**Methods:** A prospective study including 144 patients who underwent functional endoscopic sinus surgery (FESS). All participants were asked to finish SNOT-22, Lund-Mackay CT score and endoscopic examination evaluations before surgery and at 6 months after surgery. Endoscopic examination videos were evaluated using 3 scoring systems by two blinded rhinologists. The scores were compared in terms of responsiveness, validity, reliability and correlation with other scores.

**Results:** The MLK and DIP endoscopic scoring systems showed high test-retest reliability and inter-rater reliability. All endoscopic scoring systems showed significant differences between the scores recorded at each time point (all  $P < 0.05$ ). The concurrent validity of the MLK system was  $r > 0.90$  when compared with the Lund - Kennedy (LK) endoscopic scoring system and showed a statistically significant difference in discriminant validity between symptomless and symptomatic cases. The MLK and DIP systems showed high correlations with other subject assessments and no correlation with SNOT-22.

**Conclusion:** MLK and DIP exhibit substantial responsiveness, validity and reliability. MLK and DIP may be well suited for clinical and research use.

**Key words:** chronic rhinosinusitis, endoscopy, reliability, validity, correlation

## Introduction

Chronic rhinosinusitis (CRS) is a commonly reported chronic disorder that affects approximately 14% of the Chinese population<sup>(1)</sup> and is caused by chronic inflammation in the nose and sinuses lasting more than 12 weeks<sup>(2)</sup>. The condition is diagnosed based on symptoms, endoscopic examinations and CT changes. Functional endoscopic sinus surgery (FESS) has become an effective treatment for CRS that is resistant to medical management<sup>(3)</sup>. To better assess the severity of CRS, numerous scoring systems have been developed that are focussed on computed tomography (CT) findings, nasal endoscopic observations and

patient-reported outcome measures (PROMs). Endoscopic scoring systems have received less scrutiny than CT systems<sup>(4)</sup>. However, endoscopic examination clearly shows the nasal mucosa, and the images can be stored on a computer recording system<sup>(5)</sup>. This technique could provide information for diagnosis and prognosis. As therapies and techniques develop, there is a significant need for a validated endoscopic scoring system to adequately assess CRS severity.

In 1995, Lund and Kennedy proposed the Lund - Kennedy (LK) endoscopic scoring system based on polyps, oedema, discharge, scarring and crusting<sup>(6)</sup>. Although the LK system has been

reported to have a poor correlation with subjective scores <sup>(7)</sup>, it remains the most widely used system in clinical work. Some scoring systems have been developed over the last two decades. The modified Lund - Kennedy (MLK) endoscopic scoring system does not include the items of scarring and crusting <sup>(8)</sup>. The discharge, inflammation, and polyps/oedema (DIP) endoscopic scoring system was developed by removing crusting and adhesions items and increasing data points for the remaining items <sup>(9)</sup>. MLK and DIP have demonstrated good reliability. However, the validity and responsiveness of these systems remain unknown.

Therefore, the objectives of this prospective cohort study were as follows: (i) to develop a Chinese translation and cultural adaptation of the MLK and DIP based on the guidelines <sup>(10)</sup> and (ii) to validate their clinical application over a 6-month period, including the assessment of responsiveness, validity, reliability and correlation with other systems.

## Materials and methods

### Study population

A prospective study was conducted. Patients were enrolled between September 2013 and December 2016. The study was approved by the Research Ethics Committee of the First Affiliated Hospital of Wenzhou Medical University. Informed consent was obtained from the subjects, and data were collected in an anonymised database. The patients included adults 18 years and older with a diagnosis of CRS according to the American Academy of Otolaryngology–Head and Neck Surgery guidelines. Exclusion criteria included heart disease, mental disorder, malignant neoplasm of the nasal cavity and fungal sinusitis.

### Subjective assessments

The symptoms of all patients were rated using Sino-Nasal Outcome Test-22 (SNOT-22) <sup>(11)</sup> scores.

SNOT-22 is a quality-of-life assessment comprising 22 questions; each question is scored from 0 (no problem) to 5 (the problem is as bad as it can be). The overall score can range from 0 to 110, and higher scores indicate more severe quality-of-life impairment.

### Objective assessments

Each patient's CT was assessed using the Lund - Mackay score <sup>(12)</sup> before and after FESS. The Lund - Mackay score was used to assess the ostiomeatal complex and sinuses, including the maxillary, anterior ethmoidal, posterior ethmoidal, sphenoidal, and frontal sinuses. For all sinus systems, scores of 0, 1, and 2 indicate no abnormalities, partial opacification, and total opacification, respectively. For the ostiomeatal complex, scores of 0 and 2 indicate no occlusion and occlusion, respectively. The possible score for each side ranged from 0 to 12, and the total score ranged from 0 to 24. Videos, which were recorded in preoperative and

Table 1. LK endoscopic score systems.

LK system	0	1	2
Polyps	None	In middle meatus only	Beyond middle meatus
Oedema	Absent	Mild	Severe
Discharge	None	Clear and thin	Thick and purulent
Scarring	Absent	Mild	Severe
Crusting	Absent	Mild	Severe

Table 2. DIP endoscopic score systems.

DIP systems	0	5	10
Discharge	None	Thick mucus	Purulent discharge
Inflammation	None	Mild	Severe
Polyp/oedema	None	Marked oedema/ no polyps	Polyps filling nasal cavity

Table 3. MLK endoscopic score systems.

MLK system	0	1	2
Polyps	None	In middle meatus only	Beyond middle meatus
Oedema	Absent	Mild	Severe
Discharge	None	Clear and thin	Thick and purulent

postoperative endoscopic examinations, were evaluated based on the LK (Table 1), discharge, inflammation, and polyps/oedema (DIP) (Table 2) and MLK (Table 3) systems. The total scores of the LK, MLK and DIP systems are 20, 12 and 60, respectively.

### Data collection

All patients were requested to complete SNOT - 22, the Lund - Mackay score and nasal endoscopic examination before surgery. Every patient had different surgeries based on CT and endoscopic examination. Some patients underwent only uncinectomy with antrostomy and anterior ethmoidectomy. Others may have also undergone posterior ethmoidectomy, sphenoidotomy and frontal sinusotomy. In our study, patients underwent Draf type I surgery if frontal sinusotomy was necessary and the middle turbinate was preserved. All surgeries were performed by the same surgeon. After FESS, patients received standard treatment, such as intranasal corticosteroids (mometasone furoate 200 µg/day) and nasal irrigation with normal saline. All above examinations were completed again when the patients returned for a

Table 4. Subject characteristics and comorbid factors.

Characteristic	Measurement
Age, years	44.39 ± 11.43
Gender	
Male	58 (40.3%)
Female	86 (59.7%)
Unilateral CRS	32 (22.2%)
Bilateral CRS	112 (77.8%)
Polyps	98 (68.1%)
Previous surgery	20(13.9%)

follow-up after 6 months.

All videos were independently scored by two senior rhinologists according to the different endoscopic scoring systems. All raters were blinded to the patients' symptom scores and extent of surgery. Test-retest and inter-rater reliability were calculated for each system. Test-retest reliability was assessed at 1 month after the first evaluation to eliminate the impact of memory.

#### Statistical analysis

Statistical analysis was performed using SPSS 21.0 statistical software. The data are presented as the means and standard deviations (SDs).  $P < 0.05$  indicates a significant difference. The outcomes were compared between preoperative and postoperative scores using paired sample T-tests. The Spearman correlation coefficient was calculated to determine the correlations between subjective and objective assessments. The intraclass

correlation coefficient (ICC) was used to calculate inter-rater and test-retest reliabilities for all endoscopic scoring systems. The discriminant validities of the scores were confirmed using the Wilcoxon signed-rank test.

## Results

### Clinical characteristics

In total, 144 patients were enrolled in the study. The clinical characteristics of the subjects are described in Table 4. Among these patients, 86 (59.7%) were female. The mean age of the patients was  $44.39 \pm 11.43$  years. In total, 112 (77.8%) and 32 (22.2%) cases were bilateral chronic rhinosinusitis and unilateral chronic rhinosinusitis, respectively, and 98 (68.1%) and 46 (31.9%) were CRS with and without polyps, respectively. Twenty (13.9%) patients had previously undergone surgery.

### Validation

Paired sample T-tests were conducted to examine responsiveness at different times (Table 5). The mean MLK scores in preoperative patients and at 6-month follow-up were  $7.83 \pm 2.69$  and  $3.76 \pm 1.23$ , respectively; the DIP scores were  $32.93 \pm 11.27$  and  $13.85 \pm 5.74$ , respectively; and the LK scores were  $7.96 \pm 2.80$  and  $4.34 \pm 1.44$ , respectively. The t-values of LK, MLK and DIP were 15.35, 17.75 and 20.59 (all  $P < 0.05$ ), respectively, indicating a good responsiveness to nose and/or sinus surgery. Similar results were found between patients who had previously undergone surgery and those who underwent surgery for the first time. The baselines of LK, MLK and DIP were  $8.86 \pm 2.62$ ,  $8.69 \pm 2.49$  and  $36.16 \pm 10.53$ , respectively, in patients with CRSwNP. The baselines of LK, MLK and DIP were  $6.04 \pm 2.14$ ,  $6.00 \pm 2.08$  and  $26.04 \pm 9.52$ , respectively, in patients with CRSsNP

Table 5. Scores in preoperative and postoperative patients.

	preoperative	postoperative	t-value	P
SNOT22	23.18±12.63	8.19±5.00	12.02	$P < 0.05$
previous surgery	28.10±15.28	9.40±4.55	5.738	$P < 0.05$
first surgery	22.38±11.98	8.00±5.04	16.55	$P < 0.05$
LK	7.96±2.80	4.34±1.44	15.35	$P < 0.05$
previous surgery	11.10±2.57	5.30±1.53	12.40	$P < 0.05$
first surgery	7.45±2.50	4.19±1.38	14.70	$P < 0.05$
DIP	32.93±11.27	13.85±5.74	20.59	$P < 0.05$
previous surgery	40.80±13.16	19.00±7.08	10.35	$P < 0.05$
first surgery	31.66±10.40	13.01±5.03	27.67	$P < 0.05$
MLK	7.83±2.69	3.76±1.23	17.75	$P < 0.05$
previous surgery	10.20±2.63	4.50±1.47	12.73	$P < 0.05$
first surgery	7.45±2.49	3.65±1.14	23.52	$P < 0.05$

Table 6. Baseline differences between CRSwNP and CRSsNP patients.

	CRSwNP	CRSsNP	P
LK	8.86±2.62	6.04±2.14	0.045
MLK	8.69±2.49	6.00±2.08	0.05
DIP	36.16±10.53	26.04±9.52	0.441

Table 7. Reliability of Endoscopic Scoring Systems.

Inter-rater reliability	ICC	95%CI
MLK	0.68	0.57-0.76
previous surgery	0.50	0.27-0.77
first surgery	0.68	0.56-0.77
DIP	0.65	0.54-0.74
previous surgery	0.63	0.26-0.87
first surgery	0.64	0.51-0.74
LK	0.67	0.56-0.75
previous surgery	0.56	0.35-0.80
first surgery	0.68	0.56-0.77
Test-retest reliability		
MLK	0.78	0.70-0.84
previous surgery	0.81	0.56-0.92
first surgery	0.75	0.65-0.82
DIP	0.76	0.67-0.82
previous surgery	0.80	0.55-0.92
first surgery	0.73	0.64-0.81
LK	0.78	0.70-0.84
previous surgery	0.79	0.54-0.92
first surgery	0.75	0.65-0.82

(Table 6). MLK and DIP for CRSwNP patients did not significantly differ from those for CRSsNP patients ( $P > 0.05$ ), while LK showed the opposite result ( $P < 0.05$ ).

The MLK score showed the highest inter-rater reliability (inter-class coefficient [ICC]=0.68; 95% confidence interval [CI]: 0.57 to 0.76), and the same result was obtained for the in test-retest reliability over a 1-month period (ICC =0.78; 95%CI: 0.70 to 0.84). The DIP score also showed high inter-rater reliability (ICC =0.65; 95%CI: 0.54 to 0.74) and test-retest reliability (ICC =0.76; 95%CI: 0.67 to 0.82). The LK score showed similar inter-rater reliability (ICC =0.67; 95%CI: 0.56 to 0.75) and test-retest reliability (ICC =0.78; 95%CI: 0.70 to 0.84). Similar results were found between those who had undergone previous surgery and those undergoing surgery for the first time, indicating good stability

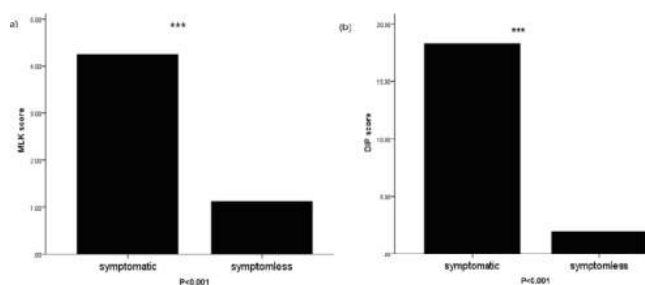


Figure 1. Discriminant validity. The MLK (a) and DIP (b) were able to discriminate rhinology clinic attendees symptomless and symptomatic side ( $n = 32$ ).  $P < 0.0001$  using the Wilcoxon signed-rank test.

regardless of previous surgery. Table 7 summarizes the reliability results obtained for each endoscopic system.

Regarding content validity, each item of the MLK and DIP scoring systems was developed depending on how the original version was developed, on literature reviews, and on discussions with experienced otolaryngologists. This thorough process aimed to help the maintenance of the practical purpose of this instrument.

To study the discriminant validity, 32 people who were diagnosed with unilateral chronic rhinosinusitis were selected. The symptomless side was used as the control group. The validity of discrimination between the symptomless and symptomatic sides was confirmed by the result of a Wilcoxon signed-rank test, which showed a statistically significant difference between the groups (Figure 1).

To study the concurrent validity, patients' response according to the LK score was compared with those according to the MLK and DIP scores at 6 months after surgery. The Spearman correlation coefficient for the MLK score was  $r = 0.92$  ( $P < 0.05$ ) and that for the DIP score was  $r = 0.73$  ( $P < 0.05$ ) (Table 8). All endoscopic scoring system scores showed high correlations with the Lund-Mackay score and no correlation with SNOT-22 in our study.

## Discussion

Several nonvalidated endoscopic scoring systems are widely used in clinics. Therefore, the aim of this study was to compare three endoscopic scoring systems in terms of their reliability, validity, responsiveness and correlation. Sinonasal endoscopic examination is important for the initial evaluation and follow-up of patients with CRS and is widely used to monitor outcomes related to various treatments during the preoperative and postoperative periods. Many researchers believe that subjective assessments do not deteriorate until periods from months to years after the nasal endoscopy findings are identified<sup>(13)</sup>. The MLK and DIP endoscopic scoring systems are objective endoscopic scores that reflect the overall inflammatory burden of disease in CRS patients. Considering that patients sensation is one of the

Table 8. Correlations between subjective and objective assessments.

	SNOT22	LK	DIP	MLK	CT
SNOT22		0.15	0.11	0.11	0.01
LK			0.73*	0.92*	0.65*
DIP				0.79*	0.57*
MLK					0.65*

main outcomes, experts suggest that CRS should be evaluated based on PROMs and objective measures<sup>(14)</sup>.

The commonly used PROMs include VAS, Sino-Nasal Outcome Test-20 (SNOT - 20)<sup>(15)</sup>, SNOT-22 and the nasal sinusitis disease index (rhinosinusitis disability index, RSDI)<sup>(16)</sup>. SNOT - 22 not only focusses on physical symptoms in patients but also considers functional limitations and emotional problems. The reliability, availability, and ease of utilization of both questionnaires were confirmed<sup>(17)</sup>.

Objective evaluation measures include Lund-Mackay CT scores, the Lund-Kennedy endoscopic scoring system, the perioperative sinus endoscopic scoring system (POSE)<sup>(18)</sup>, the DIP endoscopic scoring system and the modified Lund - Kennedy (MLK) endoscopic scoring system.

In this study, the original English versions of MLK and DIP were translated into Chinese, and a validation study was conducted using this Chinese version. When analysing the responsiveness to treatment, a significant difference was found for both the MLK and DIP scores between the initial values and those obtained at 6 months after the initial visit. This suggests that these scores are sensitive to clinical change. Interestingly, we also found that the baseline MLK and DIP scores in CRSwNP patients were higher but not significantly different from those obtained for CRSsNP patients, while the LK scores showed the opposite result. According to a reliability analysis of each endoscopic system, we found that the MLK and DIP scoring systems had high inter-rater and test-retest reliabilities regardless of whether the patients had undergone previous surgery or not, similar to results obtained in previous studies.

The content validity of the MLK and DIP scoring systems was ensured by the methodology used in developing the questionnaire. Four independent translations and two synthetic translations were obtained through a process of translation

and retranslation. We emphasized the adaptation of the local version according to international standards while maintaining the original meaning. Each of the items relates to rhinosinusitis and the array of problems associated with this condition. With regard to the concurrent validity, which indicates correlations with similar endpoints to evaluate CRS, a high correlation was observed between the MLK and LK systems. The mean values for the symptomless side were significantly lower than those for the symptomatic side in patients with unilateral chronic rhinosinusitis, suggesting that the MLK and DIP systems have fair discriminant validity.

Spearman correlation analysis of subjective and objective assessments showed that the MLK and DIP scores failed to correlate with SNOT-22, consistent with the results of other studies<sup>8</sup>. SNOT-22 has been demonstrated to be a valid disease-specific health-related QOL measure for patients with rhinosinusitis and can be affected by a variety of nasal diseases, even systemic diseases. We also found that the MLK, LK and DIP scores were highly correlated with CT score in the patients.

Ultimately, considering that the objective and subjective evaluations only provided low correlations, relying on endoscopic assessment alone cannot accurately reflect a patient's situation. EPOS 2012 proposed a staging system that is focused on a patient's reported symptoms, endoscopy and systemic medication use<sup>(19)</sup>. Thus, we suggest that a comprehensive evaluation combining subjective symptom scores and MLK and DIP scores should be used to assess a patient's condition.

## Conclusion

This study demonstrates that the MLK and DIP systems exhibit substantial responsiveness, validity and reliability. The MLK and DIP systems may be well suited for clinical and research use and might reflect rhinological conditions changes over time or the effects of an intervention.

## Authorship contribution

Lu Zhang: Organized the collection of data, performed statistical calculations and analysis, main writer of manuscript. Linghao Zhang: Planned and organized the study. Also took a major part in analysing results and writing the manuscript.

## Conflict of interest

None.

## References

1. Fan Y, Feng S, Xia W, Qu L, Li X, Chen S, et al. Aspirin-exacerbated respiratory disease in China: a cohort investigation and literature review. *Am J Rhinol Allergy*. 2012. 26(1): e20-2.
2. Benninger MS, Ferguson BJ, Hadley JA, Hamilos DL, Jacobs M, Kennedy DW, et al. Adult chronic rhinosinusitis: definitions, diagnosis, epidemiology, and pathophysiology. *Otolaryngol Head Neck Surg*. 2003. 129(3 Suppl): S1-32.
3. Georgalas C, Cornet M, Adriaansen G, Reinartz S, Holland C, Prokopakis E, et al. Evidence-based surgery for chronic rhinosinusitis with and without nasal polyps. *Curr Allergy Asthma Rep*. 2014. 14(4): 427.
4. Bhattacharyya N. The role of CT and MRI in the diagnosis of chronic rhinosinusitis. *Curr Allergy Asthma Rep*. 2010. 10(3): 171-4.
5. Murr AH, Smith TL, Hwang PH, Bhattacharyya N, Lanier BJ, Stambaugh JW, et al. Safety and efficacy of a novel bioabsorbable, steroid-eluting sinus stent. *Int*

- Forum Allergy Rhinol. 2011. 1(1): 23-32.
6. Lund VJ, Kennedy DW. Quantification for staging sinusitis. The Staging and Therapy Group. *Ann Otol Rhinol Laryngol Suppl.* 1995. 167: 17-21.
  7. Ryan WR, Ramachandra T, Hwang PH. Correlations between symptoms, nasal endoscopy, and in-office computed tomography in post-surgical chronic rhinosinusitis patients. *Laryngoscope.* 2011. 121(3): 674-8.
  8. Psaltis AJ, Li G, Vaezaefshar R, Cho KS, Hwang PH. Modification of the Lund-Kennedy endoscopic scoring system improves its reliability and correlation with patient-reported outcome measures. *Laryngoscope.* 2014. 124(10): 2216-23.
  9. Durr ML, Pletcher SD, Goldberg AN, Murr AH. A novel sinonasal endoscopy scoring system: the discharge, inflammation, and polyps/edema (DIP) score. *Int Forum Allergy Rhinol.* 2013. 3(1): 66-72.
  10. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J. Clin. Epidemiol.* 1993. 46(12):1417-32.
  11. Buckland JR, Thomas S, Harries PG. Can the Sino-nasal Outcome Test (SNOT-22) be used as a reliable outcome measure for successful septal surgery. *Clin Otolaryngol Allied Sci.* 2003. 28(1): 43-7.
  12. Lund VJ, Kennedy DW. Staging for rhinosinusitis. *Otolaryngol Head Neck Surg.* 1997. 117(3 Pt 2): S35-40.
  13. Senior BA, Kennedy DW, Tanabodee J, Kroger H, Hassab M, Lanza D. Long-term results of functional endoscopic sinus surgery. *Laryngoscope.* 1998. 108(2): 151-7.
  14. Kennedy DW, Wright ED, Goldberg AN. Objective and subjective outcomes in surgery for chronic sinusitis. *Laryngoscope.* 2000. 110(3 Pt 3): 29-31.
  15. Piccirillo JF, Merritt MG Jr, Richards ML. Psychometric and clinimetric validity of the 20-Item Sino-Nasal Outcome Test (SNOT-20). *Otolaryngol Head Neck Surg.* 2002. 126(1): 41-7.
  16. Benninger MS, Senior BA. The development of the Rhinosinusitis Disability Index. *Arch Otolaryngol Head Neck Surg.* 1997. 123(11): 1175-9.
  17. DeConde AS, Mace JC, Alt JA, Rudmik L, Soler ZM, Smith TL. Longitudinal improvement and stability of the SNOT-22 survey in the evaluation of surgical management for chronic rhinosinusitis. *Int Forum Allergy Rhinol.* 2015. 5(3): 233-9.
  18. Wright ED, Agrawal S. Impact of perioperative systemic steroids on surgical outcomes in patients with chronic rhinosinusitis with polyposis: evaluation with the novel Perioperative Sinus Endoscopy (POSE) scoring system. *Laryngoscope.* 2007. 117(11 Pt 2 Suppl 115): 1-28.
  19. Fokkens WJ, Lund VJ, Mullol J, Bachert C, Alobid I, Baroody F, et al. EPOS 2012: European position paper on rhinosinusitis and nasal polyps 2012. A summary for otorhinolaryngologists. *Rhinology.* 2012. 50(1): 1-12.

Linghao Zhang  
The First Affiliated Hospital of  
Wenzhou Medical University  
Wenzhou  
China

E-mail: 287398281@qq.com