A Comparison between W score and Ryan score in diagnosing of Laryngopharyngeal reflux disease

Gang Wang  
PLA Strategic Support Force Characteristic Medical Center

Lei Wang  
PLA Strategic Support Force Characteristic Medical Center

Zhezhe Sun  
PLA Strategic Support Force Characteristic Medical Center

Yuzhu Guo  
Beihang University

Hongdan Liu  
PLA Strategic Support Force Characteristic Medical Center

Haolun Han  
PLA Strategic Support Force Characteristic Medical Center

Changqing Zhong  
PLA Strategic Support Force Characteristic Medical Center

Xinke Sui  
PLA Strategic Support Force Characteristic Medical Center

Baowei Li  
PLA Strategic Support Force Characteristic Medical Center

Ying Zhou  
PLA Strategic Support Force Characteristic Medical Center

Changmin Qu  
PLA Strategic Support Force Characteristic Medical Center

Lianyong Li  
PLA Strategic Support Force Characteristic Medical Center

Robert T. Sataloff  
Drexel University College of Medicine

Jimin Wu  
PLA Rocket Force Characteristic Medical Center

Qi Wang  
Beijing Tongren Hospital of the Capital Medical University

Jugao Fang  
Beijing Anzhen Hospital of the Capital Medical University

Qiuping Lv  
China-Japan Friendship Hospital

Yijiang Huang  
Hainan Cancer Hospital

Wei Wu  
ent306ww@126.com

Research Article

Keywords: Laryngopharyngeal reflux, oropharyngeal pH monitoring, Ryan score

DOI: https://doi.org/10.21203/rs.3.rs-409806/v1

License: © This work is licensed under a Creative Commons Attribution 4.0 International License.  Read Full License
Abstract

Objective

To assess the diagnostic value of W score which was supposed to identify laryngopharyngeal reflux disease (LPRD) patients from the normal population by Dx-pH monitoring, comparing with Ryan score.

Methods

One hundred and eight patients with suspected LPRD and complete follow-up results after more than 8 weeks of anti-reflux therapy were enrolled from the Department of Otolaryngology-Head and Neck Surgery, Gastroenterology and Respiratory Medicine of seven hospitals. Their Dx-pH monitoring data before treatment were reanalyzed to obtain the W score besides Ryan score and then the diagnostic sensitivity and specificity were compared according to the result of anti-reflux therapy.

Results

Eighty-seven (80.6%) cases were anti-reflux therapy effective, and 21 patients (19.4%) were ineffective. Twenty-seven patients (25.0%) had a positive Ryan score. The W score was positive in 79 (73.1%) patients. There were 52 patients who had a negative Ryan score, but a positive W score. The diagnostic sensitivity, specificity, positive predictive value and negative predictive value of the Ryan score were 28.7%, 90.5%, 92.6% and 23.5%, respectively (kappa = 0.092, p = 0.068), whereas those of the W score for LPRD were 83.9%, 71.4%, 92.4% and 51.7%, respectively (kappa = 0.484, p < 0.001).

Conclusions

W score is much more sensitive for the diagnosis of LPRD. Prospective studies with larger patient populations are necessary to validate and improve the new diagnostic criteria.

Trial Registration

Chinese Clinical Trial Registry: ChiCTR1800014931.

Introduction

Laryngopharyngeal reflux disease (LPRD) is a series of diseases caused by the reflux of gastric contents up to the upper esophageal sphincter and affects as high as 10.15% of adult outpatients in the otorhinolaryngology-head and neck surgery department of class A tertiary comprehensive hospitals in China[1]. A landmark study by Koufman in 1991 demonstrated that LPRD and gastroesophageal reflux disease (GERD) should be recognized as distinct entities[2]. In contrast to typical GERD, LPRD usually does not have typical symptoms, such as acid regurgitation and heartburn, and many of its signs occur without specificity. The specificity and sensitivity of the reflux symptom index (RSI) based on reflux-related symptoms and the reflux finding score (RFS) based on laryngoscopic signs are insufficient[3]. Many experts insist that empirical drug treatment is a reliable method of LPR diagnosis [4, 5]. Nevertheless, there are many problems in the standard of starting treatment, placebo effects and patient compliance[4, 6]. It is more important to diagnose LPRD by objective means than GERD.

Presently, the most widely used objective detection method is dynamic pH monitoring. In 2009, Ayazi reported a new technique for measuring pharyngeal pH using the Restech Dx-pH Measurement System (Dx-pH) [7]. Besides its easy operation, high sensitivity and good patient tolerance, the probe can detect gas reflux and this technique has become a useful tool for diagnosing LPRD[8~12]. The Ryan score which is proposed by Dr. DeMeester has played a significant role in the diagnosis of LPRD, especially in the evaluation of operation indications and surgical efficacy for LPRD patients. Nevertheless, its sensitivity is low which means a negative Ryan score is not a sufficient indicator for excluding LPRD because of the imperfectness in its development[13]. Based on pH monitoring, the Ryan score used the normalized distances to the 95th percentage of distribution of 55 “normal” subjects who were selected merely to exclude typical GERD symptoms as the indicator. Hence, a new test with more scientific statistics is needed.

In a recent published article, we introduce a new score (named W score) based on machine learning techniques and semi-supervised leaning way[13]. In that article we elaborated the scientific and advanced nature of the test by which W score was proposed. Based on W score a subject can be diagnosed with LPRD if W > 0 and without LPRD if W < 0. Additionally, W score which benefits from the linear model structure provides an indicator of the probability that a patient is suffering from LPRD. Namely, the smaller the value is, the lower the possibility that the patient is suffering from LPRD. But whether W score performs better than Ryan score in clinical practice is still lack of strong evidence so the assessment of the efficiency was done.

Material And Methods
Study Population

This study was conducted as a retrospective analysis of patients who were suspected with LPRD and with complete 24h pharyngeal pH monitoring. All the patients came from the Department of Otorhinolaryngology, Gastroenterology and Respiratory of 6 comprehensive hospitals in China and one hospital in the U.S.A. from 2016 to March 2019. Adult patients with LPRD related symptoms (dry pharynx, pharynx itching, cough, pharynx foreign body sensation, burning sensation of the pharynx, etc.) for more than 1 month whose RSI score ≥ 10 and/or RFS score > 7 and/or Visual Analogue Scale (VAS) of the most serious symptom ≥ 5 and unsatisfactory therapeutic effect from conventional treatments were considered. All of the patients had never tried anti-reflux therapy before. One hundred and eight cases who had complete follow-up results after 2–3 months anti-reflux therapy were enrolled. The LPRD diagnosis was considered if patient responded or partially responded. The diagnosis of non-responder patients (signs and symptoms unchanged/worsened) remained uncertain and additional examinations were suggested. (Fig. 1.)

The research was conducted according to the Declaration of Helsinki and the ethics approval was obtained from the Ethics Committee of PLA 306th Hospital (K2017 [06]; Clinical Trial Registry: ChiCTR1800014931). All participants provided written informed consent.

Methodology

Scale scoring: All patients filled out RSI and underwent stroboscope prior to and after treatment. RFS was determined by two blinded senior physicians who worked independently and did not know the clinical state of the subjects in advance. VAS was scored according to the most serious symptom.

Oropharyngeal pH monitoring: Oropharyngeal pH monitoring was performed before treatment using a Dx-pH system and the Restech® pH probe (Respiratory Technology Corp., San Diego, CA, USA). After calibrating in pH 7 and pH 4 buffer solutions, the probe was inserted through the nasal cavity until its flashing LED tip was seen 5–10 mm below the uvula. The pH monitoring lasted 24 hours, and all subjects were asked to participate in normal daily activities as much as possible and to record the beginning and ending time of eating, as well as when they were in an upright and supine position.

Ryan Score was calculated from original data by software (Bio View Analysis, Sandhill Scientific, Highlands Ranch, CO, and Data View Lite, Respiratory Technology Corp, San Diego, CA). LPRD can be diagnosed if Ryan Score was > 9.41 in the upright position or > 6.79 in the supine position. Original pH data were also calculated to obtain W score and the detailed calculation methods can be found in our recent published article[13]. Based on W score a subject is diagnosed with LPRD if W > 0 and without LPRD if W < 0.

Anti-reflux therapy: Both lifestyle adjustment suggestion and anti-reflux medicine were given. According to the individual situation of each patient, specific suggestions for lifestyle adjustment were given, such as avoiding caffeine or theobromine, quitting alcohol and tobacco, increasing exercise and reducing weight. Medications were given at the same time, such as PPI, H2 antagonists, GI prokinetic agents and Gastric mucosal protective agent. Patients will be followed up after 2–3 months of anti-reflux therapy.

Data analysis

Data analysis was performed using Statistical Package for the Social Sciences, version 20.0 (SPSS, Chicago, IL, USA). The measurement data of normal distribution were expressed by mean ± SD. The data of skewed distribution were expressed by M [P25; P75] The diagnostic treatment was taken as the gold standard. A comparison of reflux related scales before and after treatment was performed with Student’s t test. Sensitivity, specificity, positive and negative predictive values, and diagnostic accuracy of the two scores were assessed respectively. The comparison of categorical data was done with Pearson Chi-squared test. P < 0.05 was considered statistically significant.

Results

One hundred and eight patients were evaluated. After 2–3 months of follow up, 87 patients (80.6%) were effective (respond + partially respond group) after anti-reflux therapy, and 21 patients (19.4%) were ineffective (non-respond group). The Ryan score was positive in 27 patients (25.0%) and W score was positive in 79 patients (73.1%). There were 52 patients who had a negative Ryan score, but a positive W score. (Table 1)
The mean RSI and RFS values of the effective group were significantly decreased after the anti-reflux therapy ($P = 0.000$), so was the VAS of the main symptom ($P = 0.000$). In the ineffective group, there was no significant difference in the mean value of RSI, VAS and RFS before and after treatment. (Table 2)

Using the clinical therapeutic effect as the standard for diagnosing LPRD, W Score had a sensitivity of 83.9% (95% CI of 74.1–90.6%), a specificity of 71.4% (95% CI of 47.7–87.8%), a positive predictive value of 92.4% (95% CI of 83.6–96.9%), and a negative predictive value of 51.7% (95% CI of 32.9–70.1%). The Ryan score had a sensitivity of 28.7% (95% CI of 19.8–39.6%) a specificity of 90.5% (95% CI of 68.2–98.3%), a positive predictive value of 92.6% (95% CI of 74.2–98.7%), and a negative predictive value of 23.5% (95% CI of 15.1–34.4%). The results showed that the sensitivity of the Ryan score was very low (28.7% v.s. 83.9%), even though it had higher specificity (90.5% v.s. 71.4%). The diagnostic efficiency of W score was better than that of Ryan score (kappa 0.484 v.s. 0.092). (Table 3)
Discussion

A series of symptoms, signs or diseases, such as dysphonia, hysteria, cough, subglottic stenosis, dysphonia, laryngeal spasm, laryngeal contact granuloma, asthma, and even chronic sinusitis and laryngeal cancer, are associated with LPRD. However, there is still a lack of a diagnostic method with high sensitivity and specificity, good patient acceptance and easy operation that has been widely recognized by clinicians. Although the RSI and RFS are proven clinical tools, they do not include many common symptoms and reflux signs and they do not take into account the frequency of symptoms[3, 14-17]. Moreover, the patients’ own emotional and psychological factors and variations in different doctors’ scores on laryngoscopy always influence the results[8]. Objective examination is needed to clarify the existence of reflux. The pH monitoring is an objective diagnostic method that accurately reflects the changes of H+ in the esophagus and/or airway. In 1969, ambulatory catheter-based esophageal pH monitoring was used to diagnose GERD[18]. In 1989, dual-probe esophageal pH monitoring was used to diagnose LPRD by Wiener et al. [19]. However, studies have shown that 45% of patients receiving dual-sensor esophageal pH monitoring have misplaced proximal sensors[20]. The Dx-pH monitoring system provides accurate probe positioning (0.5-1 cm below the uvula) and appear to be more sensitive than traditional pH monitoring in evaluation of patients with extraesophageal reflux[21].

In the analysis of 24h pH monitoring data, the Ryan Score is calculated according to the number of episodes in which pH falls below the normal range, the length of the longest episode, and the total percentage of time spent below threshold. As the pH threshold differs from the upright position and supine position, Ryan score is calculated respectively. However, the Ryan score still has some shortcomings that restrict its application. First, the Ryan score was obtained based on samples from 55 normal people[7]. The standard for normal persons was merely to exclude typical GERD symptoms such as acid reflux and heartburn, and no laryngoscopy was performed. Actually, many LPRD patients do not have gastrointestinal symptoms. Second, the pH threshold of 5.5 in the upright position and 5.0 in the supine position may be inaccurate as the pepsin is still active at pH 6.5, which can cause airway mucosal damage. In computational methods, there is no difference in the degree of discrimination between events below the threshold. Our study found that the original pH distribution data obtained by dynamic monitoring between normal people and LPRD patients are not normally distributed and that these distributions overlap[13]. However, the Ryan score obtained only from the data of normal people cannot explain the overlap of the data distribution. Although the current clinical application shows that the Ryan score can predict the efficacy of anti-reflux surgery, a negative Ryan score is still not a sufficient indicator for the exclusion of LPRD patients which means a large number of patients are misdiagnosed. Hence, a more accurate diagnostic criteria of Dx-pH monitoring for LPRD is needed.

In contrast to studies in which simple statistics were used to characterize the pH variation at the pharynx, advanced statistical methods were employed to develop W score to exploit the possibility of discriminating LPRD from normal subjects. To propose the new score, machine learning methods, which used both labelled and big unlabelled data, were employed to analyze the long-term pH data and semi-supervised learning was used to alleviate the imperfectness in data and reference test by exploring the underlying data patterns. However, it still needs further clinical verification but it is quite challenging as the lack of golden standard to diagnose LPRD. Although there are many problems in the empirical treatment, including standard of starting treatment, poor compliance, high cost, low follow-up rate and placebo effects especially, many experts insist it is a reliable and effective method of LPR diagnosis[4, 6, 22, 23]. We did this multi-center retrospective study by analyzing the data of 108 patients who underwent a formal course of anti-reflux therapy.

In our research, the inclusion criteria for anti-reflux therapy was not in accordance with the commonly used criteria (RSI > 13 and RFS > 7). Based on our previous clinical experience, patients with RSI<13 can also benefit from anti-reflux therapy. Therefore, we compared the effective rate between patients with RSI of 10-13 and patients with RSI >13, and the result showed no significant difference. In addition to RSI and RFS, we also
used VAS in combination, special for suspected LPR patients with single symptom or chief complaint not included in RSI. Although acid-suppressing therapy (PPIs and/or H2-antagonists) and GI prokinetic agents were given to ensure that patients with nonacid reflux were effectively treated without omission. Unfortunately, alginates were not given to Chinese patients due to the lack of domestic clinical drug license. Anti-reflux surgery is considered to be effective for some patients with poor medication treatment, but it had not been involved in this research [24]. The follow-up period was 2-3 months, which also reduced the false negative results caused by insufficient treatment time in some patients. We found that the sensitivity of the W score was significantly higher than that of the Ryan score, which decreased the misdiagnosis of LPRD, and the specificity of the two scores was not significantly different. More LPRD patients could be screened out through the W score and benefit from anti-reflux treatment in clinical practice.

Compared with Ryan score, the sensitivity of W index is obviously improved, but the specificity is not high enough which needs to be improved further.

Limitations

Our study has several limitations. We regarded the efficacy of anti-reflux therapy after 2-3 months as the diagnostic criterion, nevertheless, the placebo effect could not be elicited. Patients who were ineffective for anti-reflux drugs might be effective for surgery, so underestimation of diagnosis may also exist. In addition, the W score was verified only by a retrospective analysis; therefore, further validation before clinical application will be needed. At last, although alkali reflux does exist in clinical practice, W score could not pick up this kind of patient as Ryan score and more work should be done to improve it.

Conclusions

In this multi-center study, we assess the performance of a new score based on machine learning for diagnosing LPRD using the Restech Dx-pH Measurement System. The W score may be potentially useful in screening LPRD patients but needs further validation before its application in clinical practice.

Declarations

Funding:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Notation of prior abstract publication/presentation: NO

Authors' Contributions: Drs Gang Wang, Lei Wang, Zhezhe Sun had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. Drs Gang Wang, Lei Wang, Zhezhe Sun contributed equally as first authors to this work. Drs Wei Wu and Lianyong Li directed this work and contributed equally as corresponding authors.

Concept and design: Wei Wu and Lianyong Li.

Acquisition of the data: Haolun Han, Baowei Li, Ying Zhou, Xinke Sui, Robert T. Sataloff, Jimin Wu, Qi Wang, Jugao Fang, Qiuping Lv and Yijiang Huang.

Analysis, or interpretation of the data: Wei Wu, Gang Wang, Yuzhu Guo, Lianyong Li, Lei Wang, Zhezhe Sun, Haolun Han, Baowei Li, Ying Zhou and Xinke Sui.

Drafting of the manuscript: Gang Wang, Lei Wang, Zhezhe Sun.

Critical revision of the manuscript for important intellectual content: Gang Wang, Lei Wang, Zhezhe Sun, Wei Wu, Changqing Zhong and Lianyong Li.

Statistical analysis: Yuzhu Guo, Ying Zhou and Zhezhe Sun.

Administrative, technical, or material support: Hongdan Liu, Changqing Zhong, Xinke Sui.

Supervision: Wei Wu, Lianyong Li and Changmin Qu.

Conflict of Interest Disclosures: All authors declare that there are no conflicts of interest.

Additional Contributions: We thank all the LPRD patients whose participation made this study possible. None of those participants were compensated for their contribution.
References

1. Xiao, S. et al. An epidemiological survey of laryngopharyngeal reflux disease at the otorhinolaryngology–head and neck surgery clinics in China. Eur Arch Otorhinolaryngol. 277 (10), 2829–2838 (2020).

2. Koufman, J. A. The otolaryngologic manifestations of gastroesophageal reflux disease (GERD): a clinical investigation of 225 patients using ambulatory 24-hour pH monitoring and an experimental investigation of the role of acid and pepsin in the development of laryngeal injury. Laryngoscope. 101 (4 Pt 2 Suppl 53), 1–78 (1991).

3. Lechien, J. R. et al. Instruments evaluating the clinical findings of laryngopharyngeal reflux: A systematic review. Laryngoscope. 129 (3), 720–736 (2019).

4. Gupya, N., Green, R. W. & Megwalu, U. C. Evaluation of laryngopharyngeal reflux management protocol. Amer. J. Otolaryngology. 37 (3), 245–250 (2016).

5. Ford, C. N. Evaluation and management of laryngopharyngeal reflux. JAMA. 294 (12), 1534–1540 (2005).

6. Lechien, J. R. et al. Clinical outcomes of laryngopharyngeal reflux treatment: A systematic review and meta-analysis. Laryngoscope. 129 (5), 1174–1187 (2019).

7. Ayazi, S. et al. A new technique for measurement of pharyngeal pH: normal values and discriminating pH threshold. J Gastrointest Surg. 13 (8), 1422–1429 (2009).

8. Friedman, M. et al. The value of routine pH monitoring in the diagnosis and treatment of laryngopharyngeal reflux. Otolaryngol Head Neck Surg. 146 (6), 952–958 (2012).

9. Beaver, M. E. Clinical Utility of Pharyngeal pH Monitoring for Hoarseness. Otolaryngology - Head and Neck Surgery. 143(2, Supplement 2), P78(2010).

10. Williams, A. N., Simon, R. A. & Woessner, K. M. Sinusitis and chronic progressive exercise-induced cough and dyspnea. Allergy Asthma Proc. 29(6), 669–675 (2008).

11. Onyekwere, C. A., Adeyeye, O. O., Ogbera, A. O. & Duro-Emmanuel, F. Prevalence of gastroesophageal reflux disease among patients with bronchial asthma. Tropical Gastroenterology. 31 (3), 195–198 (2010).

12. Wang, C-C. et al. Airway pH monitoring in patients with suspected obstructive sleep apnoea using the Dx-pH oropharyngeal probe: Preliminary report of a prospective cohort study. Clin. Otolaryngol. 39 (6), 352–358 (2014).

13. Guo, Y. Z. et al. Machine Learning Aided Diagnosis of Diseases Without Clinical Gold Standard: A New Score for Laryngopharyngeal Reflux Disease Based on pH Monitoring. IEEE Access. 8, 67005–67014 (2020).

14. Lechien, J. R. et al. The development of new clinical instruments in laryngopharyngeal reflux disease: The international project of young otolaryngologists of the International Federation of Oto-rhino-laryngological Societies. Eur Ann Otorhinolaryngol Head Neck Dis. 135 (5), S85–S91 (2018).

15. Chang, B. A., MacNeil, S. D., Morrison, M. D. & Lee, P. K. The Reliability of the Reflux Finding Score Among General Otolaryngologists. J. Voice. 29 (5), 572–577 (2015).

16. Francis, D. O. et al. Patient-Reported Outcome Measures Related to Laryngopharyngeal Reflux: A Systematic Review of Instrument Development and Validation. Otolaryngology Head Neck Surgery. 155 (6), 923–935 (2016).

17. Kavookjian, H., Irwin, T., Garnett, J. D. & Kraft, S. The Reflux Symptom Index and Symptom Overlap in Dysphonic Patients. Laryngoscope. 130 (11), 2631–2636 (2020).

18. Spencer, J. Prolonged pH recording in the study of gastro-oesophageal reflux. Br J Surg. 56 (12), 912–914 (1969).

19. Wiener, G. et al. Chronic hoarseness secondary to gastroesophageal reflux disease: documentation with 24-h ambulatory pH monitoring. Am J Gastroenterol. 84 (12), 1503–1508 (1989).

20. McCollough, M. et al. Proximal sensor data from routine dual-sensor esophageal pH monitoring is often inaccurate. Dig Dis Sci. 49 (10), 1607–1611 (2004).

21. Yuksel, E. S. et al. An oropharyngeal pH monitoring device to evaluate patients with chronic laryngitis. Neurogastroenterol Motil. 25 (5), e315–323 (2013).

22. Gavillet, V. et al. Manifestations of gastroesophageal reflux and response to omeprazole therapy in patients with chronic posterior laryngitis: an evaluation based on clinical practice. Dig Dis Sci. 48 (11), 2117–2123 (2003).

23. Amin, M. R., Postma, G. N., Johnson, P., Diggles, N. & Koufman, J. A. Proton pump inhibitor resistance in the treatment of laryngopharyngeal reflux. Otolaryngol Head Neck Surg. 125 (4), 374–378 (2001).

24. Huang, X. et al. Endoscopic anti-reflux mucosal resection for laryngopharyngeal reflux disease: analysis of 2 cases. Chin J Dig Endosc. 38 (2), 155–156 (2021).
Figures

Figure 1
Flow chart describing the assessment of the possible LPRD patients.