Intelligence-Based Medicine 8 (2023) 100131

Contents lists available at ScienceDirect

Intelligence-Based Medicine

journal homepage: www.sciencedirect.com/journal/intelligence-based-medicine

A cascade deep learning model for diagnosing pharyngeal acid reflux episodes using hypopharyngeal multichannel intraluminal Impedance-pH signals

ARTICLE INFO

Keywords
Pharyngeal acid reflux episodes
Hypopharyngeal multichannel intraluminal impedance-pH
Laryngopharyngeal reflux
Artificial intelligence
Deep learning

ABSTRACT

Detecting pharyngeal acid reflux (PAR) episodes from 24-h ambulatory hypopharyngeal multichannel intraluminal impedance-pH (HMII-pH) signals is crucial for diagnosing laryngopharyngeal reflux (LPR). Currently, a lack of effective software for PAR episode detection requires time-consuming manual interpretation, which is prone to inter-rater variability. This study introduces a deep learning-based artificial intelligence (AI) system for PAR episode detection and diagnosis using HMII-pH signals. Ninety patients with suspected LPR and 28 healthy volunteers underwent HMII-pH testing in three Taiwanese medical centers. Candidate PAR episodes were defined as esophagopharyngeal pH drops exceeding 2 units, with nadir pH below 5 within 30 seconds during esophageal acidification. A consensus review by three experts validated 84 PAR episodes in 17 subjects. Data preprocessing identified 225 candidate PAR episodes, including 84 PAR episodes and 141 swallows/artifacts, were divided into training, validation, and test datasets (6:2:2 ratio). Three cascade deep learning AI models were trained. Among them, the cascade Multivariate Long Short-Term Memory with Fully Convolutional Network (MLSTM-FCN) model performed best in the test dataset. At the episode level, this model achieved 0.936 accuracy, 0.941 precision, 0.889 recall, 0.966 specificity, 0.914 F_1 score, and 0.864 Matthew's correlation coefficient (MCC). For subject-level evaluation, the corresponding metrics were 0.917 accuracy, 1.000 precision, 0.818 recall, 1.000 specificity, 0.900 F_1 score, and 0.842 MCC. In conclusion, the cascade MLSTM-FCN model demonstrates robust accuracy in diagnosing PAR episodes from HMII-pH signals, offering a promising tool for efficient and consistent PAR episode detection in LPR diagnosis.

1. Introduction

Laryngopharyngeal reflux (LPR) is an extraesophageal manifestation of gastroesophageal reflux disease, characterized by the backflow of stomach contents into the laryngopharynx [1]. It is a common otolaryngologic disorder associated with symptoms such as hoarseness, vocal fatigue, excessive throat clearing, globus pharyngeus, chronic cough, postnasal drip, and dysphagia. Laryngoscopic findings may include erythema, edema, ventricular obliteration, postcricoid hyperplasia, and pseudosulcus change. Non-reflux etiologies, such as voice overuse, infection, allergy, or exposure to environmental irritants, must be excluded [2]. Objective assessment of pharyngeal reflux is highly desired due to the nonspecific nature of laryngeal symptoms and signs. However, the methodology and interpretation of pharyngeal acid reflux (PAR) episodes have not yet been standardized [3]. Challenges include poor interobserver reproducibility [4] and limited outcome data associated with PAR metrics [3].

The hypopharyngeal multichannel intraluminal impedance-pH (HMII-pH) technique incorporates two *trans*-upper esophageal sphincter impedance channels to differentiate pharyngeal refluxes (retrograde impedance decreases) from swallows (antegrade impedance decreases) [5–7]. Studies have shown that the presence of one or more pharyngeal reflux episodes should be considered abnormal [8]. A composite pH parameter incorporating baseline pathological pharyngeal acid reflux (PAR) (defined as $\geq$2 PAR episodes/24 hours) and/or pathological esophageal acid exposure time has been linked to a favorable response to proton pump inhibitor (PPI) treatment in patients with suspected LPR [9]. The interobserver reproducibility of PAR episodes has also been demonstrated to be good among experienced experts [10]. However, the definition of PAR episodes is not universally accepted, and visual interpretation of PAR episodes is time-consuming and requires signal amplification [4].

Recent research by Rogers et al. has shown that machine learning (ML) decision trees (DTs) can aid in diagnosing baseline impedance measurements from pH-impedance signals [11]. However, interpreting PAR episodes based on HMII-pH signals involves multichannel output and multivariate time series processing, making DTs alone insufficient for this task. In contrast, the Multivariate Long Short Term Memory with Fully Convolutional Network (MLSTM-FCN), a novel approach for multivariate time series classification, shows promise in distinguishing PAR signals from swallow/artifact signals [12]. Additionally, a cascade ensemble deep learning artificial intelligence (AI) approach may be beneficial in handling the diverse durations and characteristics of individual PAR episodes [13]. In this study, we hypothesize that a cascade MLSTM-FCN model can effectively extract features and address time series challenges, enabling the diagnosis of PAR episodes from HMII-pH signals.

2. Materials and methods

2.1. Study design and participants

In this exploratory study, we aimed to evaluate the feasibility of using artificial intelligence (AI) for diagnosing PAR episodes based on

https://doi.org/10.1016/j.ibmed.2023.100131
Received 1 September 2023; Received in revised form 27 October 2023; Accepted 22 November 2023
Available online 25 November 2023

pH-impedance signals. The study involved 90 consecutive patients with suspected LPR, who were referred from otolaryngologic clinics at three tertiary medical centers in Taiwan: Taichung Veterans General Hospital, Chung Shan Medical University Hospital, and China Medical University Hospital. Additionally, 28 healthy volunteers were recruited through flyer distribution at Taichung Veterans General Hospital.

Before referral, each patient underwent a laryngoscopic examination and some pulmonary evaluations to exclude common upper and lower airway diseases. Subsequently, all participants underwent esophagogastroduodenoscopic examinations and HMII-pH testing when off acid suppression at the Gastrointestinal Physiological Lab of Taichung Veterans General Hospital.

The inclusion criteria consisted of the presence of major laryngeal symptoms, such as chronic cough, hoarseness, throat clearing, sore throat, and globus sensation, with at least moderate severity persisting for more than 3 months. Participants had to be 20 years of age or older. We excluded patients with chronic laryngitis caused by factors other than reflux, as previously described [9]. Healthy individuals were excluded from the study if they presented with respiratory or upper gastrointestinal symptoms, had a history of surgery, esophagitis, Barrett's esophagus, or tumors.

The enrollment period spanned from August 2016 to December 2019, and the study protocol was reviewed and approved by the Institutional Review Board of Taichung Veterans General Hospital (Approval #: CF16150B). Only de-identified data were used in this study to ensure privacy and confidentiality.

2.2. Data acquisition using HMII-pH catheters

HMII-pH catheters with 6 impedance channels and 2 pH sensors were employed to detect hypopharyngeal and esophageal reflux episodes after a minimum of 7 days off proton pump inhibitors (PPIs). The spacing between the two pH sensors (19 cm, 22 cm, or 25 cm) was determined based on individual esophageal length using catheter models CZAIBL-54, -55, and −56 (Sandhill Scientific, Inc., Highlands Ranch, CO, USA). Prior to the HMII-pH test, high-resolution manometry (HRM) was conducted using the Solar GI HRM system (MMS, Enschede, The Netherlands) to determine the positions of the upper margins of both the upper esophageal sphincter (UES) and lower esophageal sphincter (LES). The proximal pH sensor was positioned 1 cm above the manometrically determined upper margin of the UES. Consequently, the distal pH sensor was placed at a distance of 5 ± 1 cm above the upper margin of the LES [9,10]. The Bioview Analysis software (Sandhill Scientific, Highlands Ranch, CO, USA) was utilized to display the 24-h tracings, with meal times excluded.

2.3. Manual interpretation and findings of PAR episodes

A PAR episode was defined as a retrograde 50 % drop in baseline impedance starting from the more distal esophageal channel Z6 (at the level of 3 ± 1 cm above the upper margin of the LES) to the more proximal pharyngeal channel Z1 (at the level of 1 cm above the upper margin of the UES) (Fig. 1) [4]. This definition coincides with the occurrence of candidate PAR episodes in the HMII-pH tracings [10]. To ensure accuracy and avoid artifacts caused by trapped air in the pharynx, we considered only pure liquid or mixed liquid-gas PAR episodes, characterized by impedance nadirs in the pharynx of less than 1200 Ω [10,14] with a retrograde change pattern following a full column reflux of the esophagus, and without concomitant swallow episodes, as shown in Fig. 1A.

The reference standard diagnoses were determined through consensus review by three experts. Manual interpretation with experts' consensus reviews identified a total of 84 PAR episodes. Among them, 82 occurred in 16 out of 90 patients, while only 2 occurred in 1 healthy subject. There was a trend of more PAR episodes in patients compared to healthy controls, with a median (25th, 75th, 95th) number of 0 (0, 0, 3)

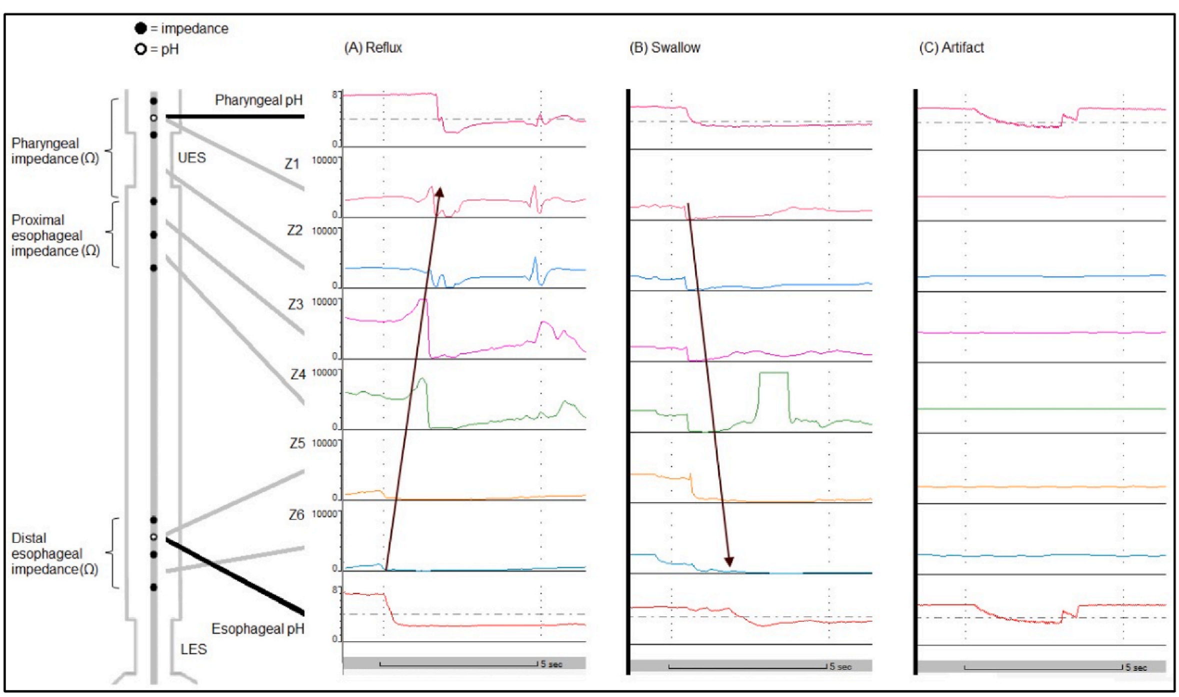

Fig. 1. Representative examples of a PAR episode, a swallow episode, and an equipment-related artifact are shown. A schematic representation depicting the recording probes and their respective placement is presented on the left-hand side. (A) A PAR episode is typically characterized by a retrograde esophagopharyngeal pH drop accompanied by a retrograde impedance decrease, starting from the more distal esophageal channel and progressing to the more proximal pharyngeal channel. (B) A swallow episode (acidic liquid swallow outside of meals) is typically characterized by an antegrade esophagopharyngeal pH drop accompanied by an antegrade impedance decrease, starting from the more proximal pharyngeal channel and progressing to the more distal esophageal channel. (C) An equipment-related artifact is typically characterized by a synchronous esophagopharyngeal drop with no change in impedance values. In this case, an abrupt return of pH to baseline was also observed. PAR, pharyngeal acid reflux.

in patients versus 0 (0, 0, 0) in healthy subjects ($P = 0.067$).

2.4. Artificial intelligence analysis of HMII-pH studies

2.4.1. Computer-aided diagnostic system

Raw data were obtained using the commercial HMII-pH analysis software Bioview (Sandhill Scientific, Highlands Ranch, CO, USA). The data were exported as comma-separated value (csv) files, excluding meal times. These files contained the multichannel raw data, which were processed using Python-based algorithms.

The computer-aided diagnostic system proposed in this study consists of two main steps: candidate PAR episode detection (data preprocessing) and PAR episode classification (Fig. 2). The detection of candidate PAR episodes is performed according to the definition outlined in section 2.4.2. Subsequently, the multichannel signals of the detected candidate PAR episodes are input into the classification model to determine the occurrence of PAR episodes, as described in section 2.4.3.

2.4.2. Data preprocessing: detection of candidate PAR episodes

Data preprocessing was conducted according to the following criteria. Candidate PAR episodes were defined as “pharyngeal pH drop of greater than 2 units and reaching a nadir pH of below 5 within 30 seconds during esophageal acidification” [10]. Previous studies by manual interpretation have demonstrated that 80 % of the proposed candidate PAR episodes correspond to HMII-pH-proven PAR episodes, and there is good inter-observer reproducibility based on experts’ consensus review [10,15,16].

To differentiate candidate PAR episodes from other events, signal magnification was utilized due to the high sampling rate of 50 Hz for both pH and impedance signals provided by the impedance-pH technology [17]. This magnification aided in excluding antegrade and synchronous esophagopharyngeal pH decreases, likely caused by swallows and artifacts, respectively (Fig. 1B & C). Additionally, common pH artifacts such as slow downward pH drift (>30 seconds to nadir pH), abrupt pH return to baseline, and out-of-range pH values (pH = 0 or > 8) were also excluded [18].

The nadir pharyngeal pH, representing the global pH minimum among local pH minima, was identified within a 30-s time window (Fig. 3). In order to detect retrograde esophagopharyngeal pH decreases, which are crucial for identifying candidate PAR episodes, the esophageal pH should ideally be < 4 at the time point of the nadir pharyngeal pH. However, in practical scenarios, the esophageal pH may occasionally be greater than 4 at the time of the nadir pharyngeal pH for a candidate PAR episode. Therefore, a 7-s time window before and after the nadir pharyngeal pH was used as a screening criterion to identify preceding esophageal acidification with pH < 4 (Fig. 4).

2.4.3. Classification of PAR episodes

2.4.3.1. Duration-specific classifiers. The durations of individual PAR episodes should be long enough to extract time series-based features from the HMII-pH signals. By combining the nadir pharyngeal pH from the candidate PAR episodes with the retrograde impedance changes in the esophagus, the duration of PAR episodes was measured from the point of 50 % reduction in baseline impedance at Z6 (the most distal esophageal impedance channel) to the nadir pharyngeal pH (Fig. 5B).

Based on the manual interpretation of PAR episodes by a consensus of three reviewers, the duration ranged from 1.1 seconds to 19.8 seconds (Fig. 5A). However, a single classifier did not yield satisfactory classification performance for such a wide range of durations. Therefore, we shortened the time period and employed multiple classifiers to cover different durations of PAR episodes, using a cascade ensemble mechanism to integrate the results.

We divided the durations of PAR episodes into three subgroups using the following criteria (Fig. 5A): Firstly, we divided 20 seconds equally into two sections, setting the cut-off point at the 10th second. Most PAR episodes occurred within the first 10 seconds. We further divided the first 10 seconds equally into two sections. As a result, the data were divided into three subgroups with cut-off points at the 5th and 10th seconds.

The time period of the first subgroup was less than or equal to 5 seconds, denoted as (0, 5]. There were 51 PAR episodes within this subgroup. Similarly, the time periods of the second and third subgroups were (5, 10] and (10, 20], comprising 28 and 5 PAR episodes, respectively.

Since the end time of candidate PAR episodes was determined by the nadir pH of the hypopharynx, we added an additional 1-s allowance after the nadir pH for feature detection. The data of each subgroup were processed by the corresponding classifier. Therefore, the time periods for the three classifiers were 6, 11, and 21 seconds (Fig. 5B).

2.4.3.2. Cascade ensemble model. Among the 118 subjects included in the study, only 17 subjects exhibited the occurrence of PAR episodes. We trained three classifiers specifically designed to diagnose PAR episodes based on three different duration lengths: 6 seconds, 11 seconds, and 21 seconds. Both PAR episodes and swallows/artifacts were divided into training, validation, and test datasets in a ratio of 6:2:2, and were fed sequentially into the three classifiers (Table 1). To avoid the possibility of data leakage, we organized the candidate PAR episodes of the same patients in the same datasets. We then tested three types of deep learning-based algorithms: Convolutional Neural Network Long Short Term Memory Network (CNN-LSTM), Convolutional Multi-Timescale Echo State Network (ConvMESN), and Multivariate Long Short Term Memory Fully Convolutional Network (MLSTM-FCN). To improve the classification performance, we applied a cascade model in individual type of algorithms, as depicted in Fig. 2.

The cascade ensemble model contained 3 sequential classifiers: 6 s,

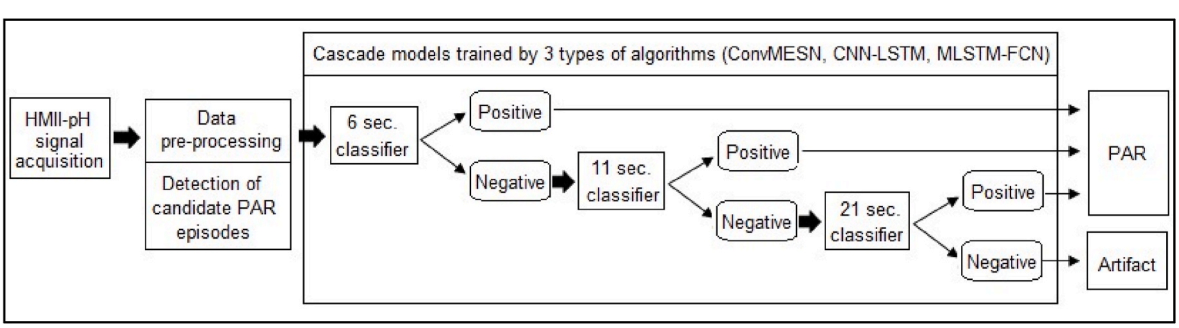

Fig. 2. Flow chart of the computer-aided diagnostic system for PAR episodes, demonstrating the data pre-processing structure and the proposed cascade ensemble model. Further details regarding the cascade ensemble model are provided in section 2.4.3. PAR, pharyngeal acid reflux.

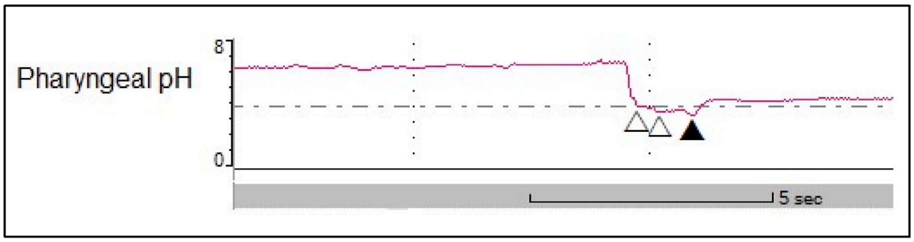

Fig. 3. Identification of the global pH minimum (indicated by the black arrowhead) as the nadir pharyngeal pH among the three local pH minima in a candidate PAR episode. PAR, pharyngeal acid reflux.

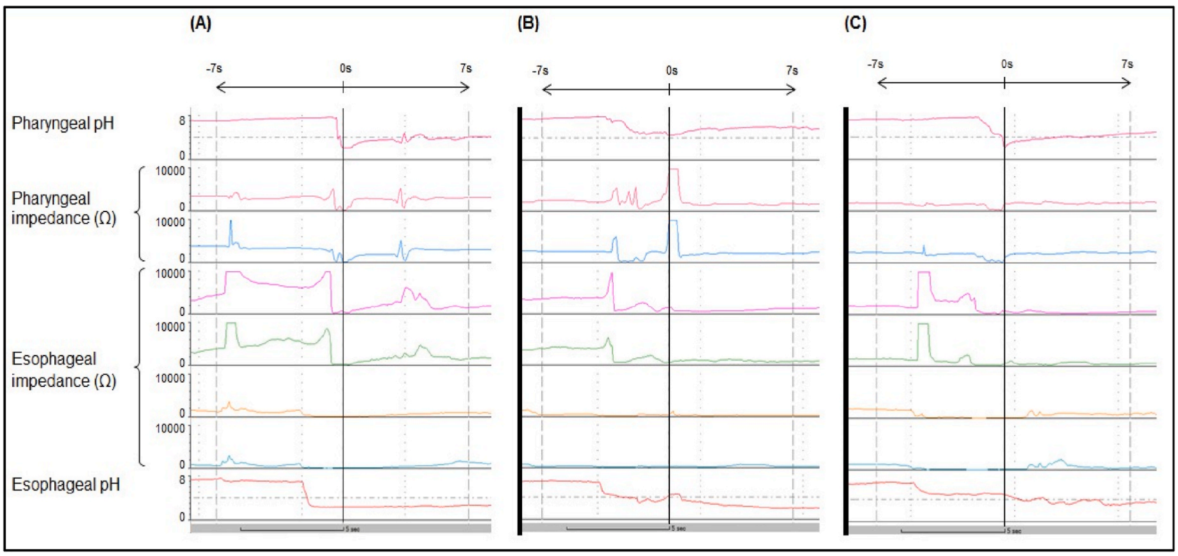

Fig. 4. Esophageal pH values at the time point of the nadir pharyngeal pH of a PAR episode. Typically, esophageal pH values are less than 4 (A), but occasionally they may be around (B) or above 4 (C). Therefore, a time window of 7 seconds before and after the nadir pharyngeal pH was utilized for the detection of preceding esophageal acidification with pH < 4. PAR, pharyngeal acid reflux.

11 s, and 21 s. The input signals were fed into the 6 s classifier trained to identify PAR episodes with a duration in the range of (0, 5] seconds. If the output was negative, the signals were then fed to the 11 s classifier trained to identify PAR episodes with a duration in the range of (5, 10] seconds. Similarly, if the output was also negative, the signals were further processed by the 21 s classifier trained to identify PAR episodes with a duration in the range of (10, 20] seconds. If the output of all three classifiers was negative, indicating no presence of PAR episodes, the input signals were classified as swallows/artifacts. On the other hand, if any of the classifiers indicated the presence of PAR episodes, the input signals were classified as PAR episodes.

2.5. *Measurement metrics*

We conducted an evaluation to assess the diagnostic performance of three distinct cascade deep-learning AI models in comparison to experts' consensus reviews for diagnosing PAR episodes. This evaluation was specifically carried out at the episode-level, focusing on the ability of the AI models to accurately identify and classify individual episodes.

For the Cascade model with the best episode-level performance, we further access the diagnostic performance in the subject-level. By this combination, we aimed to provide a more comprehensive understanding of the model's diagnostic capabilities.

To gauge the performance of the AI models, we employed several widely recognized measurement metrics. These metrics included accuracy, precision, recall (sensitivity), specificity, F_1 score, and the Mathew correlation coefficient (MCC).

3. Results

3.1. *AI extraction of candidate PAR episodes*

The supervised AI software extracted 225 candidate PAR episodes, including 84 PAR episodes and 141 swallows/artifacts (Table 1).

3.2. *The results of the test dataset in the episode-level*

The cascade MLSTM-FCN model outperformed the other two algorithms. A detailed performance comparison among the three types of models is provided in Table 2.

3.3. *The results of the test dataset in the subject-level*

The diagnostic performance of cascade MLSTM-FCN model in the subject-level is shown in Table 3.

4. Discussion

This exploratory retrospective study aimed to assess the feasibility of using AI to extract PAR (pharyngeal acid reflux) episodes from ambulatory HMII-pH tracings. We found that the supervised cascade MLSTM-FCN model provided accurate diagnoses compared to expert consensus reviews. This proof-of-concept study demonstrated that the cascade MLSTM-FCN model, which utilizes cascade ensemble deep learning AI to handle multichannel output multivariate time series data, effectively captured the features of esophageal events involving pH and impedance

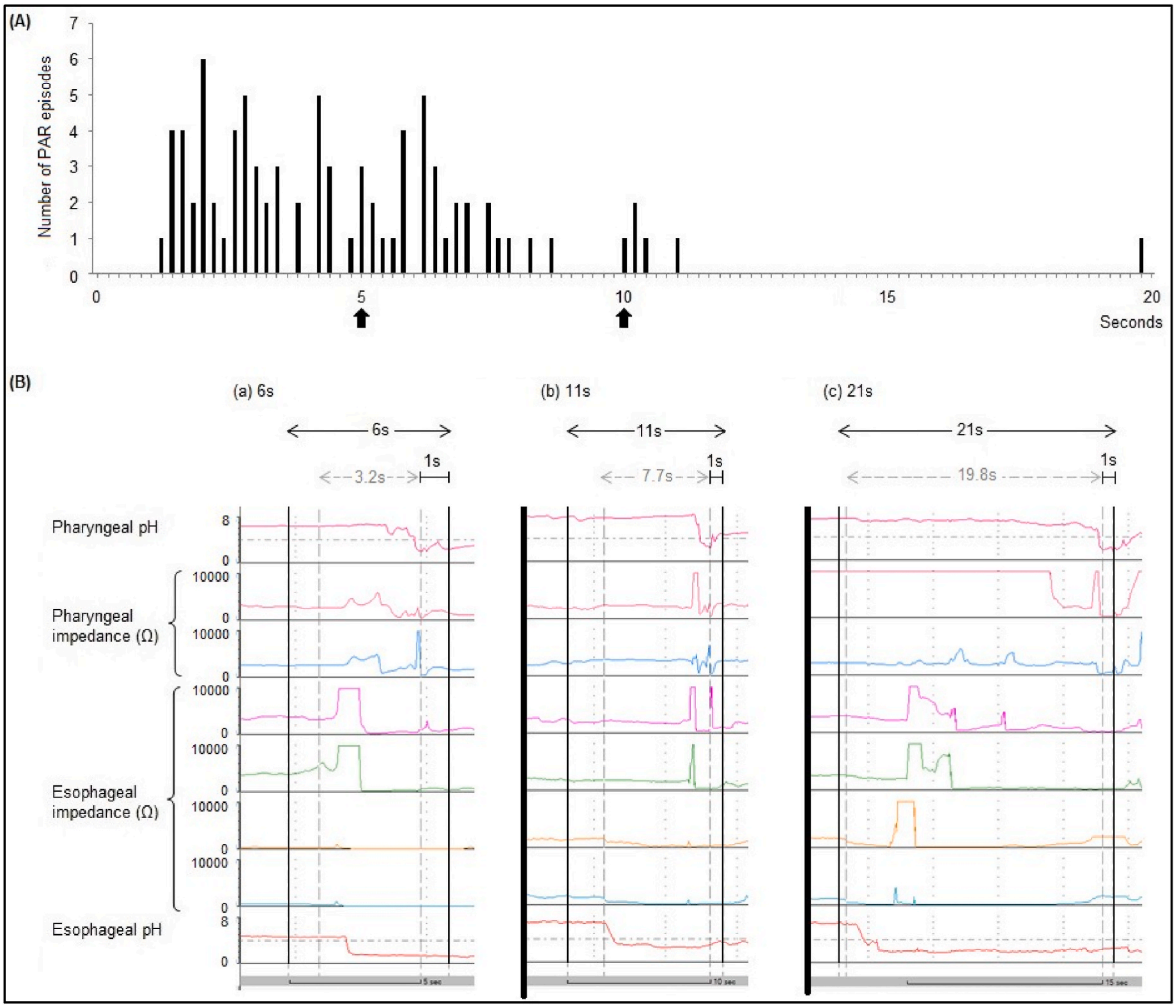

Fig. 5. (a) Histogram showing the duration of 84 PAR episodes diagnosed based on consensus review by 3 experts. The x-axis represents the duration of the PAR episodes, and the y-axis indicates the number of occurrences. Each bin on the x-axis represents an interval of 0.2 seconds. The left arrow indicates the cut-off point between the 1st and 2nd subgroups, while the right arrow indicates the cut-off point between the 2nd and 3rd subgroups. (B) Examples of PAR episodes in 3 duration-specific classifiers: (a) 6 seconds, (b) 11 seconds, and (c) 21 seconds. The duration of each PAR episode is measured between two vertical dashed lines. PAR, pharyngeal acid reflux.

Table 1
The number of PAR episodes and swallows/artifacts used to train and test 3 cascade deep learning AI models.

Type of classifier	Subgroup of data	Length of PAR episode	Model training phase				Model test phase	
			Training dataset		Validation dataset		Test dataset	
			PAR	swallows/artifacts	PAR	swallows/artifacts	PAR	swallows/artifacts
6 sec.	1st	(0, 5]	30	84*	10	28*	11	
11 sec.	2nd	(5, 10]	16	84*	6	28*	6	
21 sec.	3rd	(10, 20]	3	84*	1	28*	1	
Cascade ensemble model							18	29

*Because the cascade ensemble model consisted of 3 independently trained classifiers, data of swallows/artifacts were reused in the model training phase.
PAR, pharyngeal acid reflux; AI, artificial intelligence.

changes. While the criteria for PAR episodes are not currently available in standard impedance-pH tracing analysis, our study is the first to show that AI can objectively and reproducibly extract PAR episodes from HMII-pH tracings based on our definition.

In a previous study by Rogers et al. a complex decision tree machine learning model was used to identify 88.5 % of esophageal events in the development of AI software for interpreting MII-pH tracings [11]. However, their approach involved deleting esophageal events, including swallows and reflux events, which are characterized by impedance changes in the antegrade and retrograde directions, respectively. Additionally, their method required converting raw signals into one-dimensional feature vectors for decision tree analysis, resulting in a loss of information. In our study, we directly processed the raw multi-channel pH-impedance signals using deep learning-based AI techniques. By using the MLSTM-FCN model, which has demonstrated efficiency in complex multivariate time series classification tasks, we avoided the need for feature extraction and directly used the intact pH-impedance signals as input. However, we observed that a single classifier did not

Table 2
Diagnostic performance of 3 types of cascade deep learning models for the test dataset in the episode-level.

Model		ConvMESN		CNN-LSTM		MLSTM-FCN	
Confusion Matrix		Positive	Negative	Positive	Negative	Positive	Negative
Experts' consensus	Positive	12	6	15	3	16	2
	Negative	1	28	1	28	1	28
Accuracy		0.851		0.915		0.936	
Precision		0.923		0.938		0.941	
Recall (Sensitivity)		0.667		0.833		0.889	
Specificity		0.966		0.966		0.966	
F_1 score		0.774		0.882		0.914	
MCC		0.687		0.820		0.864	
Operation time		0:00:20.1		0:00:45.5		0:01:50.0	

ConvMESN, Convolutional Multi-timescale Echo State Network; CNN-LSTM, Convolutional Neural Network and Long Short-Term Memory Network; MLSTM-FCN, Multivariate Long Short-Term Memory Network and Fully Convolutional Network; MCC, Matthews correlation coefficient.

Table 3
Diagnostic performance of cascade MLSTM-FCN model against experts' consensus diagnosis for the test dataset in the subject-level.

		Subject-level model output	
		Positive	Negative
Experts' consensus	Positive	9	2
	Negative	0	13
Accuracy		0.917	
Precision		1.000	
Recall (Sensitivity)		0.818	
Specificity		1.000	
F_1 score		0.900	
MCC		0.842	

MLSTM-FCN, Multivariate Long Short Term Memory with Fully Convolutional Network; AI, artificial intelligence.

provide satisfactory performance for PAR episodes of various durations ranging from 1.1 s to 19.8 s. To address this, we divided the input length into three types (0–5 s, 5–10 s, 10–20 s) and trained three individual classifiers, which were then integrated into a cascade ensemble model. Experimental results showed that the cascade model provided optimal diagnostic performance.

Currently, MII-pH is considered the gold standard for diagnosing reflux episodes, as it can detect reflux and swallowing of gas, liquid, acid, and non-acid based on impedance changes and pH values of the bolus [3]. However, current automated software often overestimates non-acidic or weakly acidic reflux episodes, necessitating manual interpretations of MII-pH tracings [19,20]. Manual interpretation is time-consuming and prone to inter- and intra-rater variability, especially for pharyngeal reflux episodes [4]. In our study, we used the reviewers' consensus for manual interpretation of PAR episodes using HMII-pH tracings as the reference standard, which were recently validated by interobserver reproducibility [10]. Based on the Wingate consensus established by experts for interpreting gastroesophageal reflux episodes from MII-pH signals [21], Rogers et al. determined that acidic refluxate, high proximal extent, upright position, and longer acid clearance times were four independent characteristics with the highest concordance among experts [22], which partially explains the feasibility of our reviewers' consensus for manual interpretation of PAR episodes.

While dual pH probes have been recommended for detecting PAR episodes [2], Desjardin et al. reported that 91.3 % of simultaneous esophageal and pharyngeal pH drops below pH 5 or pH 4 regardless of decrease magnitude were due to swallows as detected by HMII-pH technology [23]. Conversely, using dual pH probes, Williams et al. showed that 78 % of rapid pharyngeal pH drops of at least 2 units to a nadir pH below 5 within 30 seconds may coincide with esophageal acidification, compared to only 8 % of pH drops of 1–2 units [24], suggesting a high possibility of true PAR based on the proposed criteria of the candidate PAR episodes used in our study. In fact, Lien et al. found that using a 3-pH-sensor and the same criteria of candidate PAR episodes, 17 % of 104 subjects with suspected laryngopharyngeal reflux had one or more candidate PAR episodes, which exhibited good inter-rater agreement [15]. Recently, they also found that 80 % of candidate PAR episodes interpreted by experts' consensus were true PAR episodes based on the HMII-pH technology, suggesting the potential clinical implication of the proposed criteria of candidate PAR episodes [10]. Although additional research is needed to confirm the reliability and validity of PAR episodes based on our criteria, the high accuracy of the supervised deep learning-based AI model for diagnosing PAR episodes in the current study corroborates its reliability.

There are limitations to this study. Firstly, it was a retrospective study with a relatively small sample size of Taiwanese participants. More studies with larger cohorts from diverse ethnic groups are needed to validate our findings. Secondly, although the MLSTM-FCN model demonstrated high diagnostic accuracy, AI-based deep learning algorithms remain black-box models, making it difficult to understand the decision-making process and identify potential confounders [25]. Thirdly, the clinical relevance and treatment outcomes specific to PAR episodes alone remain unclear.

5. Conclusion

The application of supervised deep learning-based AI, such as MLSTM-FCN, may provide sufficiently accurate diagnoses of PAR episodes using a 24-h hypopharyngeal multichannel intraluminal impedance-pH device.

Financial support

This work was funded by Taichung Veterans General Hospital, Taichung, Taiwan (TCVGH-AI-10906). This funding agency played no role in the analysis of the data or the preparation of this manuscript.

Author contributions

Han-Chung Lien and Jachih Fu conceived and designed the experiments; Ping-Huan Lee, Chen-Chi Wang, Ying-Cheng Lin, Chun-Yi Chuang, Yung-An Tsou, Yen-Yang Chen, Sheng-Shun Yang and Han-Chung Lien performed the experiments; Jachih Fu, Ping-Huan Lee and Han-Chung Lien analyzed the data; and Jachih Fu, Ping-Huan Lee and Han-Chung Lien wrote the manuscript and approval of the final version.

Declaration of competing interest

None for all the authors (Jachih Fu, Ping-Huan Lee, Chen-Chi Wang, Ying-Cheng Lin, Chun-Yi Chuang, Yung-An Tsou, Yen-Yang Chen, Sheng-Shun Yang, Han-Chung Lien).

Acknowledgements

We deeply appreciate Miss Kareen Chong for her secretarial work. We deeply appreciate Min-Xiu Lian and Chia-Wei Lin, National Yunlin University of Science and Technology graduate students, helped with the programming and data analysis.

References

[1] Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R, Global Consensus Group. The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. Am J Gastroenterol 2006;101:1900–20.

[2] Koufman JA, Aviv JE, Casiano RR, Shaw GY. Laryngopharyngeal reflux: position statement of the committee on speech, voice, and swallowing disorders of the American Academy of Otolaryngology-Head and Neck Surgery. Otolaryngol Head Neck Surg 2002;127:32–5.

[3] Gyawali CP, Kahrilas PJ, Savarino E, Zerbib F, Mion F, Smout AJPM, et al. Modern diagnosis of GERD: the lyon consensus. Gut 2018;67:1351–62.

[4] Zerbib F, Roman S, Bruley Des Varannes S, Gourcerol G, Coffin B, Ropert A, et al. Normal values of pharyngeal and esophageal 24-hour pH impedance in individuals on and off therapy and interobserver reproducibility. Clin Gastroenterol Hepatol 2013;11:366–72.

[5] Borges LF, Chan WW, Carroll TL. Dual pH probes without proximal esophageal and pharyngeal impedance may be deficient in diagnosing LPR. J Voice 2019;33: 697–703.

[6] Suzuki T, Seki Y, Okamoto Y, Hoppo T. Hypopharyngeal multichannel intraluminal impedance leads to the promising outcome of antireflux surgery in Japanese population with laryngopharyngeal reflux symptoms. Surg Endosc 2018;32: 2409–19.

[7] Hou C, Chen M, Chen T, Yang Y, Yang X, Lin Z, et al. Study on laryngopharyngeal and esophageal reflux characteristics using 24-h multichannel intraluminal impedance-pH monitoring in healthy volunteers. Eur Arch Oto-Rhino-Laryngol 2020;277:2801–11.

[8] Hoppo T, Sanz AF, Nason KS, Carroll TL, Rosen C, Normolle DP, et al. How much pharyngeal exposure is "normal"? Normative data for laryngopharyngeal reflux events using hypopharyngeal multichannel intraluminal impedance (HMII). J Gastrointest Surg 2012;16:16–24.

[9] Lien HC, Wang CC, Kao JY, Yeh HZ, Hsu JY, Lee SW, et al. Distinct physiological characteristics of isolated laryngopharyngeal reflux symptoms. Clin Gastroenterol Hepatol 2020;18:1466. 74.e4.

[10] Chen YY, Wang CC, Lin YC, Kao JY, Chuang CY, Tsou YA, et al. Validation of pharyngeal acid reflux episodes using hypopharyngeal multichannel intraluminal impedance-pH. J Neurogastroenterol Motil 2023;29:49–57.

[11] Rogers B, Samanta S, Ghobadi K, Patel A, Savarino E, Roman S, et al. Artificial intelligence automates and augments baseline impedance measurements from pH-impedance studies in gastroesophageal reflux disease. J Gastroenterol 2021;56: 34–41.

[12] Karim F, Majumdar S, Darabi H, Harford S. Multivariate LSTM-FCNs for time series classification. Neural Network 2019;116:237–45.

[13] Ambalavanan KA, Devarakonda MV. Using the contextual language model BERT for multi-criteria classification of scientific articles. J Biomed Inf 2020;112: 103578.

[14] Kawamura O, Aslam M, Rittmann T, Hofmann C, Shaker R. Physical and pH properties of gastroesophagopharyngeal refluxate: a 24-hour simultaneous ambulatory impedance and pH monitoring study. Am J Gastroenterol 2004;99:1000–10.

[15] Lien HC, Wang CC, Hsu JY, Sung FC, Cheng KF, Liang WM, et al. Classical reflux symptoms, hiatus hernia, and overweight are independent risk factors of pharyngeal acid reflux in patients with suspected reflux laryngitis. Aliment Pharmacol Ther 2011;33:89–98.

[16] Lien HC, Wang CC, Liang WM, Sung FC, Hsu JY, Yeh HZ, et al. Composite pH predicts Esomeprazole response in laryngopharyngeal reflux without typical reflux syndrome. Laryngoscope 2013;123:1483–9.

[17] Pritchett JM, Aslam M, Slaughter JC, Ness RM, Garrett CG, Vaezi MF. Efficacy of esophageal impedance/pH monitoring in patients with refractory gastroesophageal reflux disease, on and off therapy. Clin Gastroenterol Hepatol 2009;7:743–8.

[18] Harrell SP, Koopman J, Woosley S, Wo JM. Exclusion of pH artifacts is essential for hypopharyngeal pH monitoring. Laryngoscope 2007;117:470–4.

[19] Roman S, Gyawali CP, Savarino E, Yadlapati R, Zerbib F, Wu J, et al. Ambulatory reflux monitoring for diagnosis of gastro-esophageal reflux disease: update of the Porto consensus and recommendations from an international consensus group. Neuro Gastroenterol Motil 2017;29:1–15.

[20] Kang HJ, Park JM, Choi SY, Kim SI, Lee YC, Eun YG, et al. Comparison between manual and automated analyses in multichannel intraluminal impedance: pH monitoring for laryngopharyngeal reflux. Otolaryngol Head Neck Surg 2022;166: 128–32.

[21] Gyawali CP, Rogers B, Frazzoni M, Savarino E, Roman S, Sifrim D. Inter-reviewer variability in interpretation of pH-impedance studies: the wingate consensus. Clin Gastroenterol Hepatol 2021;19:1976. 8.e1.

[22] Rogers B, Rogers B, Frazzoni M, Savarino E, Roman S, Sifrim D, et al. Episode-level reflux characteristics: how 6 experienced reviewers differentiate true reflux from artifact on pH-impedance studies. Neuro Gastroenterol Motil 2022;34:e14153.

[23] Desjardin M, Roman S, des Varannes SB, Gourcerol G, Coffin B, Ropert A, et al. Pharyngeal pH alone is not reliable for the detection of pharyngeal reflux events: a study with oesophageal and pharyngeal pH-impedance monitoring. United European Gastroenterol J 2013;1:438–44.

[24] Williams RB, Ali GN, Wallace KL, Wilson JS, De Carle DJ, Cook IJ. Esophagopharyngeal acid regurgitation: dual pH monitoring criteria for its detection and insights into mechanisms. Gastroenterology 1999;117:1051–61.

[25] Le Berre C, Sandborn WJ, Aridhi S, Devignes MD, Fournier L, Smaïl-Tabbone M, et al. Application of artificial intelligence to gastroenterology and hepatology. Gastroenterology 2020;158:76–94.e2.

Jachih Fu[1]
Computer Aided Measurement and Diagnostic Systems Laboratory, Department of Industrial Engineering and Management, National Yunlin University of Science and Technology, Yunlin, Taiwan

Ping-Huan Lee[1]
Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan

Chen-Chi Wang
Department of Otolaryngology, Taichung Veterans General Hospital, Taichung, Taiwan
School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan
School of Speech Language Pathology & Audiology, Chung Shan Medical University, Taichung, Taiwan

Ying-Cheng Lin
Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan

Chun-Yi Chuang
Department of Otolaryngology, Chung Shan Medical University Hospital, Taichung, Taiwan
School of Medicine, Chung Shan Medical University, Taichung, Taiwan

Yung-An Tsou
Department of Otolaryngology-Head and Neck Surgery, China Medical University Hospital, Taichung, Taiwan

Yen-Yang Chen
Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan

Sheng-Shun Yang
Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan
School of Medicine, Chung Shan Medical University, Taichung, Taiwan
Institute of Biomedical Sciences, National Chung Hsing University, Taichung, Taiwan

Han-Chung Lien*
Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan
School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan
Department of Post-Baccalaureate Medicine, College of Medicine, National Chung Hsing University, Taichung, Taiwan

* Corresponding author. Department of Post-Baccalaureate Medicine, College of Medicine, National Chung Hsing University, Taichung, Taiwan School of Medicine, National Yang Ming Chiao Tung University, Taipei, Taiwan Division of Gastroenterology, Taichung Veterans General Hospital, Taichung, Taiwan 1650, Boulevard Sect. 4, Taichung 40705, Taiwan.
E-mail address: lhc@vghtc.gov.tw (H.-C. Lien).

[1] Jachih Fu and Ping-Huan Lee contributed equally to this work as first authors.