Voice Enabled Artificial Intelligence for Detection of Pathologic Gastroesophageal Reflux Disease and Barrett's Esophagus

Amrit K. Kamboj, MD[1], Manoj Yarlagadda, MBBS[1], Mary Pietrowicz, PhD[2], Kevin Buller[1], Prasad G. Iyer, MD, MS[1], David A. Katzka, MD[3], Keiko Ishikawa, PhD[4], Diana Orbelo, PhD[5], Cadman L. Leggett, MD[1]

[1]Division of Gastroenterology and Hepatology, Mayo Clinic, Rochester, MN; [2]Applied Research Institute, University of Illinois at Urbana-Champaign, Champaign, IL; [3]Division of Gastroenterology and Hepatology, Columbia University, New York City, NY; [4]Speech and Hearing Science, University of Illinois, Urbana-Champaign, IL; [5]Department of Otolaryngology, Mayo Clinic, Rochester, MN

BACKGROUND

- Gastroesophageal reflux disease (GERD) can lead to voice alterations, including hoarseness.
- The study aim was to identify specific voice biomarkers associated with pathologic GERD using advanced machine learning tools.
- Voice biomarkers reflect the periodicity of the voice signal, which mirrors the quality of the voice. The lesser the periodicity, the more noisy the sound quality (e.g., hoarseness, breathiness, roughness, creakiness).
- Detection of pathologic GERD, including Barrett's esophagus (BE), with voice biomarkers can serve as a simple non-invasive screening tool.

STUDY SUBJECTS

- Voice recordings were obtained from patients undergoing clinically indicated esophagogastroduodenoscopy (EGD) and/or ambulatory pH monitoring studies.
- Patients were excluded if they had another condition (pulmonary, cardiac, neurologic, etc) associated with voice disturbance.
- Voice recording consisted of a 6-sentence standard script read over 25-45 seconds.

STUDY GROUPS

- GERD(+) patients were defined as those with erosive esophagitis (LA grade B-D) or peptic stricture or acid exposure time >6%.
- BE was defined as columnar mucosa >1 cm with confirmed specialized intestinal metaplasia.
- Patients without these findings were considered GERD(-).
- A vocally normal group consisting of individuals with normal voice as judged by speech pathology evaluation was used as an independent control group.

STATISTICAL ANALYSIS

- Random forest models were trained using a balanced number of subjects per condition using random participant selection from the majority class.
- Using a 5-fold nested cross validation strategy, features were selected and ranked within fold, and a series of models were trained within each fold using recursive feature elimination.
- The average F1 score, a harmonic mean of precision and recall (range 0-100), across all folds was reported to assess performance.

RESULTS

- The study sample consisted of 245 patients (vocally normal, n=98; GERD(-), n=78; GERD(+), n=34; BE, n=35).
- Feature rankings suggested voice quality differences between groups relating to voice signal periodicity.
- The model demonstrated excellent ability to discern BE and GERD(+) from the vocally normal group with F1 scores 82 (males) and 89 (females) and 80 (males) and 80 (females) for BE and GERD(+) respectively.
- There was also a good voice signal distinguishing BE and GERD(+) groups from GERD(-) with F1 scores ranging from 60-70.

BASELINE CHARACTERISTICS

Subgroup	Sex	Number of patients	Mean age, years (SD)
Barrett's esophagus	Female	13	65 (15)
	Male	22	67 (9)
GERD (+)	Female	23	51 (14)
	Male	11	53 (19)
GERD (-)	Female	48	55 (15)
	Male	30	59 (15)
Vocally normal	Female	64	28 (11)
	Male	34	34 (15)

Table 1. Baseline patient characteristics in study subgroups.

RECEIVER OPERATING CHARACTERISTICS

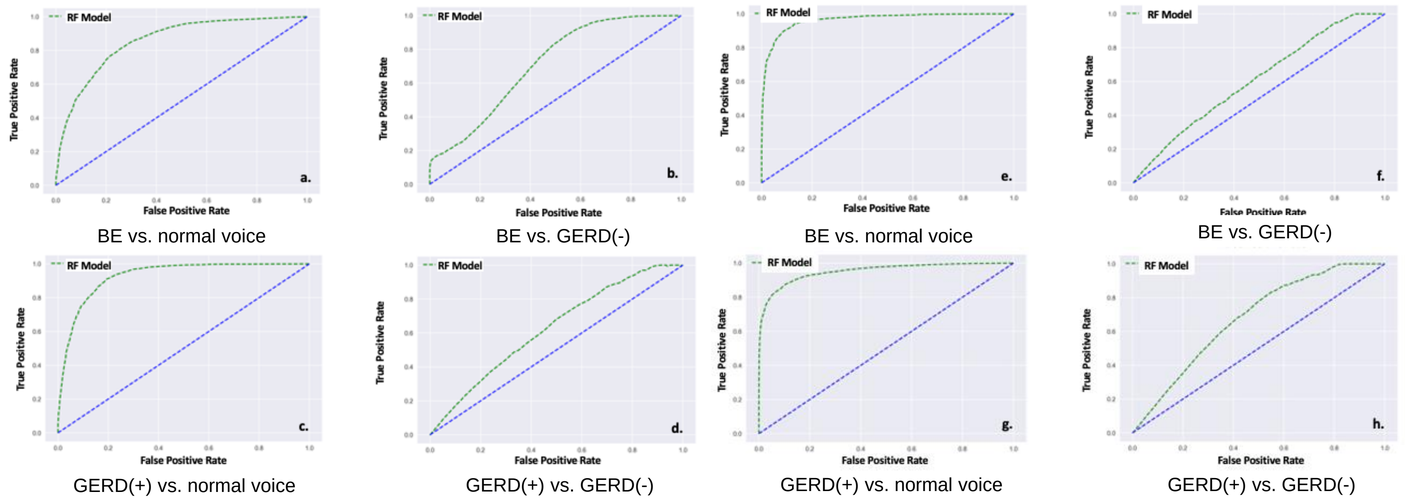

Figure 1. Receiver Operating Characteristics for a Random Forest (RF) Model for males (a-d) and females (e-h).

DISCUSSION

- These results suggest that voice biomarkers may be useful as a non-invasive tool in the detection of pathologic GERD/BE.
- A deep learning diagnostic model will be developed using the identified voice biomarkers.

©2022Mayo Foundation for Medical Education and Research