



Published in final edited form as:

Neurogastroenterol Motil. 2021 October ; 33(10): e14106. doi:10.1111/nmo.14106.

Diagnostic Yield and Reliability of Post-Prandial High-Resolution Manometry and Impedance-pH for Detecting Rumination and Supragastric Belching in PPI Non-Responders:

Physiologic Testing for Behavioral Disorders

Kelli DeLay¹, John E. Pandolfino², Sabine Roman³, C. Prakash Gyawali⁴, Edoardo Savarino⁵, Michael Tye², Alexander Kaizer¹, Rena Yadlapati^{1,6}

¹Division of Gastroenterology and Hepatology, University of Colorado Anschutz Medical Campus, Aurora, Colorado, USA;

²Division of Gastroenterology and Hepatology, Department of Medicine, Feinberg School of Medicine, Northwestern University, Chicago, Illinois, USA;

³Université de Lyon, Hospices Civils de Lyon, Digestive Physiology, Hopital E Herriot, Lyon, France; Université de Lyon, Lyon I University, Digestive Physiology, Lyon, France; Université de Lyon, Inserm U1032, LabTAU, Lyon, France;

⁴Division of Gastroenterology, Washington University School of Medicine, St. Louis, Missouri, USA;

⁵Department of Surgery, Oncology and Gastroenterology, Division of Gastroenterology, University of Padua, Padova, Italy;

⁶UCSD Center for Esophageal Diseases, Division of Gastroenterology, University of California San Diego, San Diego, California; USA.

Abstract

Background: Supragastric belching (SGB) and rumination are behavioral disorders associated with proton pump inhibitor (PPI) non-response and can be diagnosed using multichannel intraluminal impedance pH (MII-pH) and post-prandial high-resolution impedance manometry (PPHRIM). This pilot study compared diagnostic yield and inter-rater agreement for SGB and rumination using MII-pH and PPHRIM.

Author and Address for Correspondence: Kelli DeLay, MD, Assistant Professor of Clinical Medicine, Division of Gastroenterology & Hepatology, University of Colorado Anschutz Medical Campus, 12631 E. 17th Ave. MS B-158, Aurora, CO 80045, 720-723-6779, kelli.delay@CUAnschutz.edu.

Disclosures and Conflicts of Interest

John E. Pandolfino: Consultant for Medtronic, Croston, Diversatek, Torax, Ironwood, Impleo. Pharmaceuticals

S Roman: Consultant for Medtronic, Diversatek

C. Prakash Gyawali: Consultant for Medtronic, Diversatek, Ironwood

Edoardo Savarino: Consultant for Medtronic, Reckitt Benckiser, Takeda, Merck & Co, Bristol-Myers Squibb, Abbvie, Amgen, Novartis, Fresenius Kabi, Sandoz, Sofar, Malesci, Janssen, Grifols, Aurora Pharma, Innovamedica, Johnson&Johnson, SILA, Unifarco, Alfasigma, Shire, EG Stada Group

Rena Yadlapati: Consultant: Medtronic, Ironwood Pharmaceuticals, Diversatek; Research support: Ironwood Pharmaceuticals; Advisory Board: Phathom Pharmaceuticals, RJS Mediadiagnostix

Kelli DeLay: No competing interests

Methods: Three esophageal physiologists performed blinded interpretations of MII-pH and PPHRIM in 22 PPI non-responders. Raters selected from 4 diagnostic impressions (normal, GERD, behavioral disorders, GERD+behavioral disorders) without clinical context. Primary outcomes were diagnostic impressions compared against clinical gold standard impression, between raters, and between test modalities. Following a 28-month wash-out period, raters re-interpreted MII-pH with clinical context and under consensus definition of diagnostic criteria.

Key Results: Compared to gold standard, rater accuracy for presence of behavioral disorders ranged from 45–77% on MII-pH and 45–59% on PPHRIM. On MII-pH, inter-rater agreement was fair for diagnosis (κ 0.32, $p < 0.01$) and suboptimal for presence of behavioral disorders (κ 0.13, $p = 0.14$). On PPHRIM, inter-rater agreement was suboptimal for both diagnosis (κ 0.03, $p = 0.34$) and presence of a behavioral disorder (κ -0.22, $p = 0.96$). Inter-rater agreement improved in *post-hoc* MII-pH interpretations. Rumination was more frequently identified on PPHRIM (23, 35%) compared to MII-pH (7, 11%).

Conclusions and Inferences: Diagnostic accuracy and inter-rater agreement are higher for MII-pH than PPHRIM, and behavioral disorders are more frequently identified on PPHRIM. Identifying behavioral disorders on MII-pH and PPHRIM has implications for clinical evaluation of PPI non-response; clinical context is essential for accurate study interpretation. Further work is needed to standardize definitions and interpretations.

Keywords

Gastro-esophageal reflux disease; Behavioral disorders; Belching disorders

INTRODUCTION

Proton pump inhibitor (PPI) therapy is the mainstay for treatment of suspected gastroesophageal reflux disease (GERD), however, up to 50% of patients will not derive an adequate symptom response.^{1, 2} While refractory pathologic GERD may explain PPI non-response in some patients, other common etiologies include functional upper gastrointestinal (GI) disorders, esophageal motility disorders, and importantly, behavioral disorders such as supragastric belching (SGB) or rumination syndrome.^{3, 4} A recent retrospective study of 542 patients undergoing evaluation for anti-reflux surgery for PPI non-response detected excessive SGB or rumination in over 10% of patients.⁵ With SGB and rumination syndrome increasingly recognized as mimickers of PPI-refractory GERD, it is important to accurately identify these entities so that diagnosis-driven treatment plans can be pursued, and invasive or other unnecessary interventions are minimized.^{3, 6}

Esophageal physiologic testing utilized to differentiate behavioral disorders from GERD include multichannel intraluminal impedance-pH (MII-pH) monitoring and esophageal high-resolution impedance manometry (HRIM).⁷ MII-pH has specifically been used to study patients with excessive belching, and has been shown to identify SGB as a source of symptoms.^{10, 11, 13} Distinct from MII-pH, postprandial HRIM (PPHRIM) is applied after standard HRIM to identify SGB and rumination by evaluating manometric responses to meal ingestion.^{8–10}

Despite this potential, both MII-pH and PPHRIM have limited penetrance outside of tertiary centers as tools to evaluate for behavioral disorders. Although societal recommendations advocate for esophageal physiologic testing to assess for SGB and/or rumination syndrome among PPI non-responders, the optimal strategy in clinical practice is poorly delineated.^{11, 12} This study aimed to compare the diagnostic yield of SGB and/or rumination and assess inter-rater agreement between and within PPHRIM and MII-pH studies performed in a cohort of PPI non-responders.

METHODS

Study Design

This prospective blinded pilot study was performed between February 2017 and March 2020. This study protocol was approved with a waiver of consent by the institutional review board at Northwestern University for the purpose of determination of inter-rater agreement in identification of mechanisms of symptom generation in PPI non-responders, utilizing de-identified data with no direct links to protected health information.

MII-pH and PPHRIM Studies

MII-pH and PPHRIM studies in PPI non-responders were identified from the esophageal database at Northwestern University (Chicago, Illinois, USA). Studies were selected if both a PPHRIM and MII-pH study off PPI were performed in a PPI non-responder within a one-week interval between January 2010 to May 2016. PPI non-responders were defined as patients with persistent typical reflux symptoms of heartburn and/or regurgitation despite double dose PPI therapy for 8 weeks.¹³ Each study was de-identified by the study team to remove patient identifiers and to disassociate the PPHRIM study from the MII-pH study.

All PPHRIM studies were performed using a 36-channel solid-state catheter (Manoscan 360; Medtronic, Minneapolis, Minnesota), utilizing a standardized HRIM protocol described previously.⁹ Following the standardized HRIM protocol, patients ingested a solid refluxogenic test meal (self-identified by the patient as typically symptom inducing), and postprandial monitoring was performed up to 90 minutes. Studies were excluded if the HRIM study did not include at least seven technically adequate high-quality wet swallows in the supine position or if the duration of postprandial monitoring was less than 20 minutes. Esophageal MII-pH monitoring was performed off PPI using an ambulatory multi-channel intraluminal impedance and pH monitoring system (Bioview, Diversatek, Highlands Ranch, Colorado), using a protocol outlined in previous studies.^{14, 15} Patients recorded their meals, activities, and symptoms during the 24-hour recording period.

Gold Standard

The gold standard diagnostic impression was provided by the investigator from Northwestern University (JEP), who did not participate in assessment of inter-rater agreement. This gold standard diagnostic impression was the clinical diagnosis determined in the context of clinical history and with both PPHRIM and MII-pH studies available for comparison of findings.

The following definitions were used for the gold standard analysis based on international guidelines and consensus recommendations. On MII-pH, abnormal reflux burden was defined as acid exposure time (AET) greater than 6% and/or greater than 80 reflux episodes in 24 hours.¹⁶ Reflux hypersensitivity was defined as normal reflux burden with positive symptom association: symptom index (SI) > 50 and symptom association probability (SAP) > 95%.^{17, 18} SGB was characterized by a rapid aboral movement of gas greater than or equal to 1000 Ω followed by a quick expulsion of gas in retrograde direction with a return to baseline impedance.¹⁰ Rumination was diagnosed when liquid reflux, classified by a retrograde 50% decrease in impedance from baseline, reached the proximal esophagus, largely exhibited after eating and comprised of nonacidic, food-buffered gastric contents.^{19, 20} For PPHRIM, SGB were diagnosed in the presence of relaxation of the upper esophageal sphincter (UES) and anterograde air flow identified by a rapid rise in impedance by 1000 Ω followed by rapid expulsion of air with diaphragm and esophagogastric junction (EGJ) contraction, UES relaxation, and/or negative intrathoracic pressure.^{9, 10, 21} Rumination was identified when abdominal pressure increased > 30 mmHg with a retrograde pressure gradient and flow of gastric refluxate in the absence of retching, detected by a 50% decrease in impedance from baseline, reaching the proximal esophagus up to 10 seconds after the gastric strain.^{10, 19, 22} While there are no established HRIM parameters to diagnose GERD, the presence of reflux-associated transient lower esophageal sphincter relaxations (tLESR), characterized by gastric distention followed by prolonged lower esophageal sphincter (LES) relaxation in the absence of a swallow, and/or morphologic or motor abnormalities of the EGJ or esophageal body, have been classically associated with GERD and were used to diagnose GERD on HRIM.^{8, 23}

Raters

Three raters external to the study team (SR, CPG, ES) were invited to participate given 1) international recognition as experts in esophageal physiology, 2) contributions to modern-day interpretation of MII-pH and PPHRIM, and 3) diverse representation.^{8, 9, 11, 24–26} The three raters comprised a mean of 20 years of experience and represented esophageal motility practices in the United States, France, and Italy.

Raters were instructed to interpret each study and provide their diagnostic impression categorized into two general categories: Behavioral (any SGB and/or rumination component), and Non-Behavioral (Normal, Abnormal reflux burden, and Reflux hypersensitivity; all without any behavioral component) (Table 1).

Note that reflux hypersensitivity was not an option for PPHRIM studies. For MII-pH studies raters were also instructed to indicate the total esophageal acid exposure time (AET; % time pH < 4.0), number of reflux events, symptom index, and symptom association probability. Raters were not provided a clinical history or clinical context at this point of the study.

Definitions for diagnostic impression were not discussed prior to rater interpretation such as the minimum number of symptoms for a diagnosis of reflux hypersensitivity, the number of SGB or rumination episodes for a behavioral diagnosis, or the requirement that symptom index and symptom association probability scores are both > 50 and >95, respectively, for a diagnosis of reflux hypersensitivity.

Post-hoc Study Design

Following the *a priori* analysis, a consensus meeting was held between raters and study team. Investigators agreed to conduct a post-hoc analysis to simulate pragmatic MII-pH testing interpretation and evaluate the impact of applying both clinical context and definitions for diagnostic impression, described above, on diagnostic accuracy. In this post-hoc analysis, the same 22 MII-pH were resent to raters, under different computer-generated file names, after a 28-month wash-out period after the original analysis to ensure sufficient time lapse between the two interpretations. Clinical summaries for each of the 22 patients were provided to each rater. Standard metrics for interpretation agreed upon during the consensus meeting included: 1) Keeping pH drops despite prolonged drifts, 2) Avoiding deleting presumed meal periods, even if not marked by patients, 3) Meeting four criteria for the diagnosis of reflux hypersensitivity: Normal reflux burden, three or more of each symptom must be present, SI must be > 50%, and SAP must be > 95%, and 4) Requiring that 13 episodes of supragastric belches be present to diagnose SGB.²⁷

Outcomes

The primary outcome was diagnostic impression compared against the gold standard, between blinded raters, and between test modality. Diagnostic impression included four diagnostic groups (Normal, GERD, GERD and Behavioral, Behavioral), used for the gold standard interpretation and the three expert raters. Behavioral impression consisted of whether there was any behavioral component (Behavioral or No Behavioral component present) (Table 1).

Secondary outcomes included categorical acid exposure time (physiologic (< 4.0%), inconclusive (4.0 to 6.0%), and pathologic (> 6.0%)), and continuous measurements of symptom index and symptom association probability.

Data Analysis

Categorical summary measures are presented as count (percent). Agreement was measured by percent overall agreement and either Cohen's kappa or Fleiss' kappa to account for contexts with two or three "raters", respectively, to account for the categorical ratings. Kappa values from 0.81–1.00 were considered almost perfect agreement, 0.61–0.80 substantial agreement, 0.41–0.60 moderate agreement, 0.21–0.40 fair agreement, 0.01–0.20 slight agreement, and < 0.0 poor agreement. Comparisons of proportions used a two-sample test of proportions. All analyses were completed in R v3.6.3 (Vienna, Austria).

RESULTS

Gold Standard:

Of 22 MII-pH studies performed off PPI therapy, 11 (50%) were classified as behavioral (4 GERD+behavioral, 7 behavioral) and 11 (50%) were non-behavioral (6 Normal, 5 GERD). Of 22 PPHRIM studies, 12 (55%) were behavioral (10 GERD+behavioral, 2 behavioral) and 10 (45%) were non-behavioral (3 Normal, 7 GERD). Of the 11 MII-pH studies classified as behavioral, SGB was identified in 10 of 11 (91%) and rumination in 4 of 11 studies (36%). For PPHRIM, SGB was diagnosed in 11 of 12 behavioral studies (92%) and rumination

in 4 of 12 (33%). There was fair agreement between MII-pH and PPHRIM studies for group diagnosis (45% agreement; κ 0.3 (95% CI 0.04, 0.56), $p < 0.01$) and almost perfect agreement for presence of a behavioral component (95% agreement; κ 0.9 (95% CI 0.73, 1.0), $p < 0.01$).

MII-pH Studies: Rater Accuracy and Agreement

Compared to the gold standard, rater accuracy for whether a behavioral disorder was present ranged from 45% to 77% (Figure 1B). Of the 27/66 studies raters identified as having a behavioral disorder, SGB was diagnosed in 20 (74%) and rumination 2 (7%). Inter-rater agreement for group diagnosis was fair (κ 0.32 (95% CI 0.17, 0.47), $p < 0.01$) and slight for presence of a behavioral diagnosis (κ 0.13 (95% CI -0.11, 0.37), $p = 0.14$) (Table 2). Inter-rater agreement was highest for GERD+behavioral diagnosis (κ 0.52 (95% CI 0.08, 0.82), $p < 0.01$). Inter-rater agreement for AET category was fair (κ 0.36 (95% CI 0.16, 0.55), $p < 0.01$), and unanimous for SI and SAP (κ 1.00, $p < 0.01$).

Post-hoc Inter-Rater Agreement for MII-pH Studies: Following the consensus meeting between raters and the study team establishing guidelines for MII-pH interpretation, the re-named 22 MII-pH studies with clinical summaries were dispersed to the raters for repeat interpretation. Agreement improved across all measures. Inter-rater agreement for diagnostic group was moderate (κ 0.47 (95% CI 0.30, 0.64); $p < 0.01$) with agreement for diagnostic group across all three raters for 12 (55%) MII-pH studies. (Table 2). Inter-rater agreement for presence of behavioral diagnoses was slight (κ 0.19 (95% CI -0.07, 0.45); $p = 0.16$). Among 22 interpretations per rater for a total of 66 interpretations, 24/66 changed in post-hoc interpretation. Of 38 initially interpreted as behavioral, 20 changed to non-behavioral. Of 28 initially interpreted as non-behavioral, 4 changed to behavioral. Inter-rater agreement for AET category was near perfect (κ 0.93 (95% CI 0.72, 1.0); $p < 0.01$).

PPHRIM Studies: Rater Accuracy and Agreement

Compared to the gold standard, rater accuracy for whether a behavioral condition was present ranged 45 to 59% (Figure 1B). Of the 46/66 studies raters identified as having a behavioral disorder, SGB was diagnosed in 23 (50%) and rumination in 20 (43%). Inter-rater agreement for diagnostic group was slight (κ 0.03 (95% CI -0.12, 0.18), $p = 0.34$) and poor for presence of a behavioral diagnosis (κ -0.22 (95% CI -0.46, 0.02), $p = 0.96$).

Inter-Test Agreement Between MII-pH and PPHRIM:

Rater individual analysis of MII-pH and PPHRIM studies agreed in 54% of cases for presence of a behavioral diagnosis (κ 0.03 (95% CI -0.21, 0.27), $p = 0.39$). Behavioral diagnosis (SGB and/or rumination) was more frequently identified on PPHRIM than MII-pH (46/66 (70%) vs 38/66 (58%); $p = 0.20$) (Table 3). In particular, SGB was identified in 25/66 (38%) interpretations for MII-pH and in 26/66 (39%) interpretations for PPHRIM, and rumination was identified in 7/66 (11%) MII-pH interpretations and 23/66 (35%) for PPHRIM.

DISCUSSION

This study was undertaken to address the substantial need to understand the clinical role of MII-pH and PPHRIM in identification of behavioral diagnoses such as SGB and rumination syndrome among PPI non-responders. Our study of three blinded raters prospectively interpreting 22 MII-pH and 22 PPHRIM studies from PPI non-responders demonstrated that 1) inter-rater agreement on MII-pH for presence of a behavioral diagnosis is higher compared to PPHRIM, 2) raters agree on presence of a behavioral diagnosis on MII-pH when clinical context and standardized definitions are available to the rater, and 3) PPHRIM more frequently identifies presence of a behavioral diagnosis, particularly rumination episodes, compared to MII-pH. These findings highlight the clinical value of MII-pH and PPHRIM in the evaluation of a PPI non-responder, particularly in diagnosing behavioral disorders using established criteria, provided an accurate history and clinical context are available.

In this study the gold standard was the diagnostic impression from an expert esophageal physiologist who interpreted both MII-pH and PPHRIM studies in each patient with clinical context. Gold standard agreement for whether a behavioral disorder was present in both MII-pH and PPHRIM studies was nearly perfect (95% agreement), whereas the rater agreement was substantially lower without clinical context (54% agreement). These findings highlight the importance of clinical context for study analysis and necessity of establishing guidelines for interpretation, including methodology to address isolated pH drops and meal periods. When this information was provided, the inter-rater agreement for presence of a behavioral condition improved from slight to fair, and categorical AET improved from fair to near perfect.

PPHRIM more often identified a behavioral event, particularly rumination, and this may be explained by its unique technical attributes. With simultaneous assessment of pressure changes from the UES to the intra-abdominal cavity, and directionality of liquid and gas bolus flow, HRIM can identify distinct pressure and bolus flow patterns, particularly rumination;^{10, 21} in fact, raters identified rumination in 43% of PPHRIM cases compared to 2% of MII-pH cases where a behavioral disorder was identified. Because of the absence of pressure data on MII-pH, changes in impedance with or without a drop in pH could result in rumination falsely identified as a reflux episode⁹ (Figure 2a). In a prospective study of patients with symptoms of post-prandial regurgitation undergoing MII-pH, AET was similar between patients with regurgitation-predominant refractory GERD and rumination syndrome, which further highlights the potential for misdiagnosis.²⁸ On the other hand, since SGB identification relies on characteristic impedance changes without need for pressure data, MII-pH may outperform PPHRIM in diagnosing SGB, especially since impedance changes and patient symptom reporting can be monitored for 24 hours with MII-pH.²⁹ Raters diagnosed SGB in 74% of MII-pH studies where a behavioral disorder was present compared to 50% on PPHRIM.

While PPHRIM may more effectively identify rumination, the inter-rater agreement in this study was sub-optimal for PPHRIM. There may be several reasons contributing to this disagreement. Compared to MII-pH, PPHRIM is a newer concept that has primarily been

adopted by tertiary care centers. The protocol for performing and interpreting PPHRIM is not standardized, and manometric software does not yet automate identification of rumination and belch patterns. Further, despite the widespread use of coupling impedance to identify retrograde bolus movement with TLESR to suggest GERD, there is still no standard definition of GERD with HRIM. The fact that diagnosis of behavioral disorders improved among raters after incorporation of clinical context and diagnostic criteria suggests that development of specific parameters for identification of SGB and rumination on PPHRIM will improve the value of this test.

There are important limitations to this study. This is a pilot study, and the cohort studied is relatively small, therefore a comparison of testing modalities is potentially underpowered, affecting our kappa statistics. Multiple diagnostic impressions may also have overestimated agreement. The combined expertise of the three raters likely exceeds that of most providers, and thus the results may overestimate diagnostic agreement for behavioral conditions such as SGB and rumination and limit generalization of our findings. The process itself of a post-hoc analysis after discovering only slight agreement for a behavioral diagnosis from initial MII-pH interpretation may have biased raters towards a behavioral process when re-reviewing these studies. While a post-hoc analysis was conducted to assess inter-rater agreement for MII-pH after clinical context was provided, the same post-hoc process was not done for PPHRIM, and the expert rater (JEP) providing the gold standard diagnosis did not participate in the post-hoc analysis to evaluate internal consistency. Additionally, we did not assess agreement for PPHRM versus PPHRIM, and the refluxogenic meals incorporated into PPHRIM were not standardized. There are no established thresholds for diagnosing GERD on HRIM, however diagnostic criteria could include 1) identification of tLESR with associated reflux during relaxation,³⁰ 2) evaluation of mean baseline impedance, found to have high diagnostic accuracy for GERD and significantly lower in GERD than controls (1061 vs 2814 Ω ; $p < .0001$),³¹ 3) structural components of EGJ integrity i.e. presence of a hiatal hernia and EGJ-contractile integral,^{11,32} and/or 4) motor abnormalities: the presence of esophageal hypomotility or motor fragmentation.³³ The absence of increased intragastric pressure preceding refluxate helps to exclude rumination and narrow the diagnosis to GERD. Finally, to distinguish GERD from SGB or rumination on HRiM, it has been proposed that SGB could be defined as > 2 episodes/hour with/without tLESRs and rumination syndrome occurring at least once/hour, with/without tLESRs or SGB.⁹

Though there was potential for rater recall of prior MII-pH interpretation, the MII-pH studies were re-labeled and the analysis delayed for 28 months to circumvent possible recall. While novel pH-impedance parameters including post-reflux swallow-induced peristaltic wave (PSPW) index and mean nocturnal baseline impedance (MNBI) have been used as adjuncts in diagnosing refractory GERD, we did not incorporate these testing into analysis to reflect a more practical clinical evaluation across testing centers.³⁴ It is important to note that this study should not serve as an estimation of prevalence of rumination and/or belching disorders as the 22 studies were performed at a tertiary referral academic center for esophageal diseases with pretest suspicion of a behavioral condition. Further, the technology available at this academic center may not be widely accessible among other academic institutions or community practices.

In conclusion, despite these limitations, our study provides significant findings and importantly, highlights directions for future research. Our results support the clinical role of MII-pH off PPI therapy and PPHRIM in the evaluation of PPI non-responders. Distinguishing between PPI-refractory GERD and behavioral disorders is critically important in avoiding inappropriate procedures and shifting towards vastly different therapeutic interventions including cognitive-behavioral therapy, baclofen, or neuromodulation. MII-pH may have higher accuracy and inter-rater agreement in general and for SGB, and PPHRIM should be considered when rumination is suspected. Assessing competence for providers specializing in esophageal motility is a priority, and these data support the importance of teaching to identify behavioral conditions with both MII-pH and PPHRIM studies. Optimally, manometric and MII-pH software should automate identification of belch patterns and rumination. With regards to PPHRIM, the protocol for performing and interpreting the study needs to be standardized, as has been done in pediatrics.^{35, 36} Certainly, outcomes and clinical impact of utilizing esophageal physiologic testing for PPI non-responders should be examined through chosen treatment plans, patient outcomes, and health care utilization and costs. Though further study is needed, these data support that in patients where a behavioral condition is highly presumed, both MII-PH and PPHRIM should be considered and, if performed, interpreted by an experienced provider competent in study interpretation, with knowledge of diagnostic definitions and clinical context.

Research Funding Support:

This study was funded by NIH R01 DK092217-04 (PI: Pandolfino)

Data available on request from the authors: The data that support the findings of this study are available from the corresponding author, KD, upon reasonable request.

Abbreviations:

AET	acid exposure time
CI	confidence interval
EGJ	esophagogastric junction
GERD	gastroesophageal reflux disease
GI	gastrointestinal
HRIM	high-resolution impedance manometry
LES	lower esophageal sphincter
MII-pH	multichannel intraluminal impedance-pH
MNBI	mean nocturnal baseline impedance
PPHRM	post-prandial high-resolution manometry
PPHRIM	post-prandial high-resolution impedance manometry

PPI	proton pump inhibitor
PSPW	post-reflux swallow-induced peristaltic wave
SAP	symptom association probability
SGB	supragastric belching
SI	symptom index
TLESR	transient lower esophageal sphincter relaxation

References

1. Delshad SD, Almario CV, Chey WD, et al. Prevalence of Gastroesophageal Reflux Disease and Proton Pump Inhibitor-Refractory Symptoms. *Gastroenterology* 2020;158:1250–1261.e2. [PubMed: 31866243]
2. Kahrilas PJ, Boeckstaens G, Smout AJ. Management of the patient with incomplete response to PPI therapy. *Best Pract Res Clin Gastroenterol* 2013;27:401–14. [PubMed: 23998978]
3. Halland M, Pandolfino J, Barba E. Diagnosis and Treatment of Rumination Syndrome. *Clin Gastroenterol Hepatol* 2018;16:1549–1555. [PubMed: 29902642]
4. Ferron GM, McKeand W, Mayer PR. Pharmacodynamic Modeling of Pantoprazole's Irreversible Effect on Gastric Acid Secretion in Humans and Rats. *The Journal of Clinical Pharmacology* 2001;41:149–156. [PubMed: 11210394]
5. Sawada A, Guzman M, Nikaki K, et al. Identification of Different Phenotypes of Esophageal Reflux Hypersensitivity and Implications for Treatment. *Clin Gastroenterol Hepatol* 2020.
6. Roman S, Mion F. Refractory GERD, beyond proton pump inhibitors. *Curr Opin Pharmacol* 2018;43:99–103. [PubMed: 30240968]
7. Garbarino S, Horton A, Patel A. The Utility of Esophageal Motility Testing in Gastroesophageal Reflux Disease (GERD). *Curr Gastroenterol Rep* 2019;21:37. [PubMed: 31289922]
8. Roman S, Holloway R, Keller J, et al. Validation of criteria for the definition of transient lower esophageal sphincter relaxations using high-resolution manometry. *Neurogastroenterol Motil* 2017;29.
9. Yadlapati R, Tye M, Roman S, et al. Postprandial High-Resolution Impedance Manometry Identifies Mechanisms of Nonresponse to Proton Pump Inhibitors. *Clin Gastroenterol Hepatol* 2018;16:211–218.e1. [PubMed: 28911949]
10. Kessing BF, Bredenoord AJ, Smout AJ. Objective manometric criteria for the rumination syndrome. *Am J Gastroenterol* 2014;109:52–9. [PubMed: 24366235]
11. Gyawali CP, Kahrilas PJ, Savarino E, et al. Modern diagnosis of GERD: the Lyon Consensus. *Gut* 2018.
12. Yadlapati R, Pandolfino JE. Personalized Approach in the Work-up and Management of Gastroesophageal Reflux Disease. *Gastrointest Endosc Clin N Am* 2020;30:227–238. [PubMed: 32146943]
13. Yadlapati R, Vaezi MF, Vela MF, et al. Management options for patients with GERD and persistent symptoms on proton pump inhibitors: recommendations from an expert panel. *Am J Gastroenterol* 2018.
14. Savarino E, Zentilin P, Tutuian R, et al. The role of nonacid reflux in NERD: lessons learned from impedance-pH monitoring in 150 patients off therapy. *Am J Gastroenterol* 2008;103:2685–93. [PubMed: 18775017]
15. Zerbib F, Roman S, Ropert A, et al. Esophageal pH-impedance monitoring and symptom analysis in GERD: a study in patients off and on therapy. *Am J Gastroenterol* 2006;101:1956–63. [PubMed: 16848801]

16. Roman S, Gyawali CP, Savarino E, et al. Ambulatory reflux monitoring for diagnosis of gastro-esophageal reflux disease: Update of the Porto consensus and recommendations from an international consensus group. *Neurogastroenterol Motil* 2017;29:1–15.
17. Bredenoord AJ, Weusten BL, Timmer R, et al. Addition of esophageal impedance monitoring to pH monitoring increases the yield of symptom association analysis in patients off PPI therapy. *Am J Gastroenterol* 2006;101:453–9. [PubMed: 16464226]
18. Weusten BL, Roelofs JM, Akkermans LM, et al. The symptom-association probability: an improved method for symptom analysis of 24-hour esophageal pH data. *Gastroenterology* 1994;107:1741–5. [PubMed: 7958686]
19. Kessing BF, Govaert F, Masclee AA, et al. Impedance measurements and high-resolution manometry help to better define rumination episodes. *Scand J Gastroenterol* 2011;46:1310–5. [PubMed: 21815865]
20. Saleh CM, Bredenoord AJ. Utilization of esophageal function testing for the diagnosis of the rumination syndrome and belching disorders. *Gastrointest Endosc Clin N Am* 2014;24:633–42. [PubMed: 25216908]
21. Kessing BF, Bredenoord AJ, Smout AJ. Mechanisms of gastric and supragastric belching: a study using concurrent high-resolution manometry and impedance monitoring. *Neurogastroenterol Motil* 2012;24:e573–9. [PubMed: 23072402]
22. Absah I, Rishi A, Talley NJ, et al. Rumination syndrome: pathophysiology, diagnosis, and treatment. *Neurogastroenterol Motil* 2017;29.
23. Ho SC, Chang CS, Wu CY, et al. Ineffective esophageal motility is a primary motility disorder in gastroesophageal reflux disease. *Dig Dis Sci* 2002;47:652–6. [PubMed: 11911355]
24. Savarino E, Tutuian R, Zentilin P, et al. Characteristics of reflux episodes and symptom association in patients with erosive esophagitis and nonerosive reflux disease: study using combined impedance-pH off therapy. *Am J Gastroenterol* 2010;105:1053–61. [PubMed: 19997095]
25. Savarino E, Marabotto E, Zentilin P, et al. The added value of impedance-pH monitoring to Rome III criteria in distinguishing functional heartburn from non-erosive reflux disease. *Dig Liver Dis* 2011;43:542–7. [PubMed: 21376679]
26. Gyawali CP, Fass R. Management of Gastroesophageal Reflux Disease. *Gastroenterology* 2018;154:302–318. [PubMed: 28827081]
27. Koukias N, Woodland P, Yazaki E, et al. Supragastric Belching: Prevalence and Association With Gastroesophageal Reflux Disease and Esophageal Hypomotility. *J Neurogastroenterol Motil* 2015;21:398–403. [PubMed: 26130635]
28. Nakagawa K, Sawada A, Hoshikawa Y, et al. Persistent Postprandial Regurgitation vs Rumination in Patients With Refractory Gastroesophageal Reflux Disease Symptoms: Identification of a Distinct Rumination Pattern Using Ambulatory Impedance-pH Monitoring. *Am J Gastroenterol* 2019;114:1248–1255. [PubMed: 31246694]
29. Kessing BF, Bredenoord AJ, Smout AJ. The pathophysiology, diagnosis and treatment of excessive belching symptoms. *Am J Gastroenterol* 2014;109:1196–203; (Quiz) 1204. [PubMed: 25001253]
30. Ribolsi M, Holloway RH, Emerenziani S, et al. Impedance-high resolution manometry analysis of patients with nonerosive reflux disease. *Clin Gastroenterol Hepatol* 2014;12:52–7. [PubMed: 23891920]
31. Ravi K, Geno DM, Vela MF, et al. Baseline impedance measured during high-resolution esophageal impedance manometry reliably discriminates GERD patients. *Neurogastroenterol Motil* 2017;29.
32. Gyawali CP, Roman S, Bredenoord AJ, et al. Classification of esophageal motor findings in gastro-esophageal reflux disease: Conclusions from an international consensus group. *Neurogastroenterol Motil* 2017;29.
33. Savarino E, Gemignani L, Pohl D, et al. Oesophageal motility and bolus transit abnormalities increase in parallel with the severity of gastro-oesophageal reflux disease. *Aliment Pharmacol Ther* 2011;34:476–86. [PubMed: 21671968]
34. Frazzoni M, Frazzoni L, Tolone S, et al. Lack of improvement of impaired chemical clearance characterizes PPI-refractory reflux-related heartburn. *Am J Gastroenterol* 2018;113:670–676. [PubMed: 29681623]

35. Chial HJ, Camilleri M, Williams DE, et al. Rumination syndrome in children and adolescents: diagnosis, treatment, and prognosis. *Pediatrics* 2003;111:158–62. [PubMed: 12509570]
36. Hartmann AS, Poulain T, Vogel M, et al. Prevalence of pica and rumination behaviors in German children aged 7–14 and their associations with feeding, eating, and general psychopathology: a population-based study. *Eur Child Adolesc Psychiatry* 2018;27:1499–1508. [PubMed: 29675593]

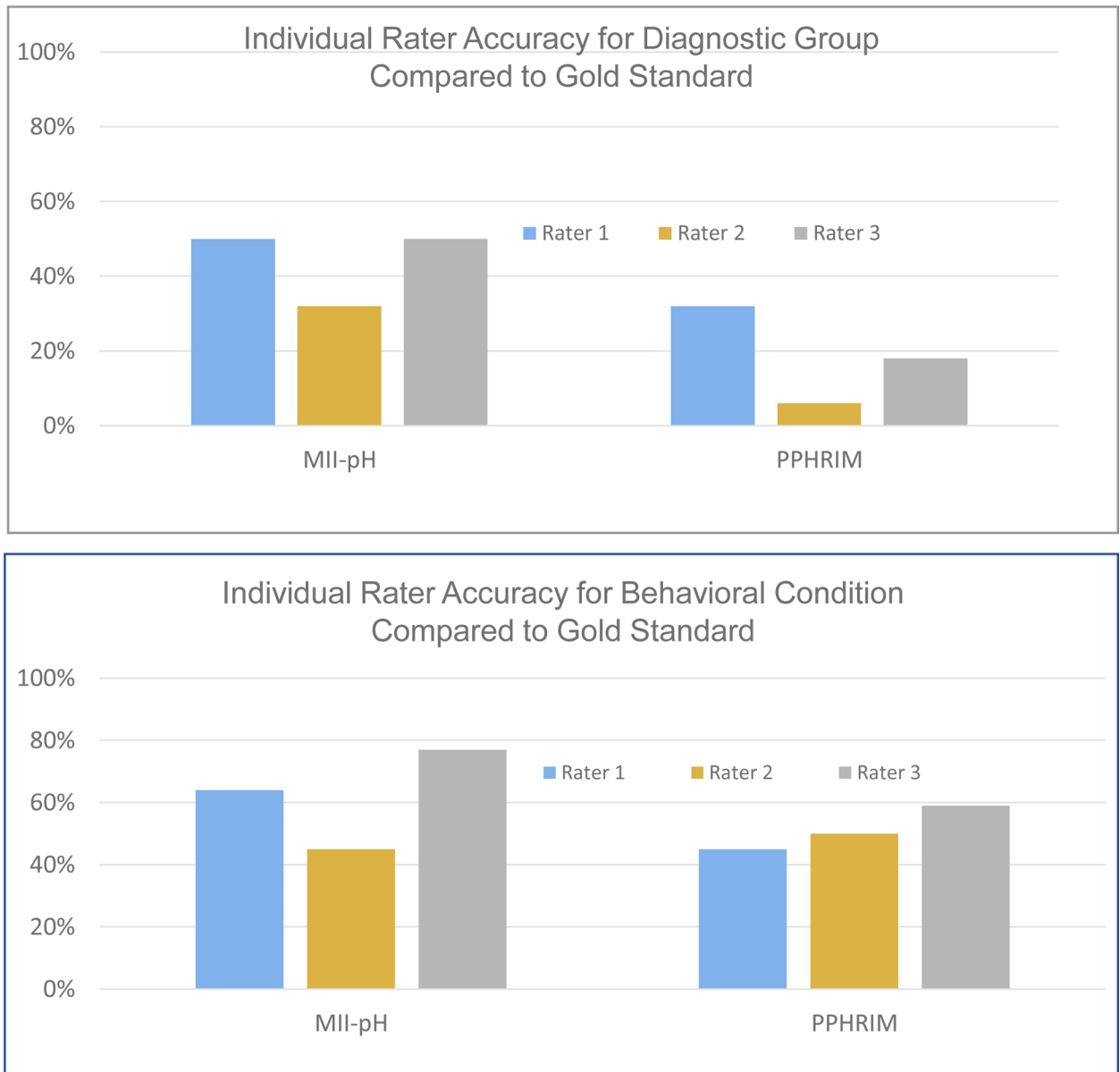
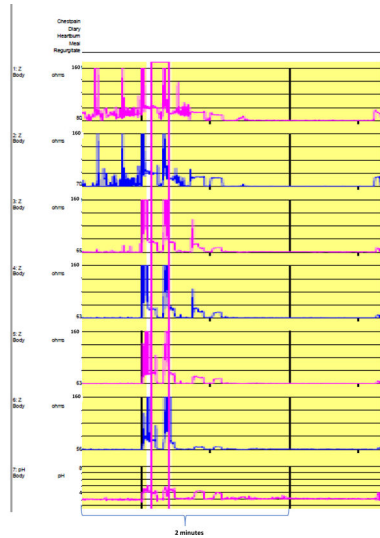


Figure 1. Individual Rater Accuracy Compared to Gold Standard for Diagnostic Group and Behavioral Condition.

(a) Rater accuracy for diagnostic group was higher in MII-pH ranging from 32% to 50% compared to 6–32% for PPHRIM. (b) Rater accuracy was similar across testing modalities for identifying a behavioral condition at 45%–77% for MII-pH and 45%–59% for PPHRIM. Rater accuracy for MII-pH includes findings from the phase 1 analysis only.

2A



2B

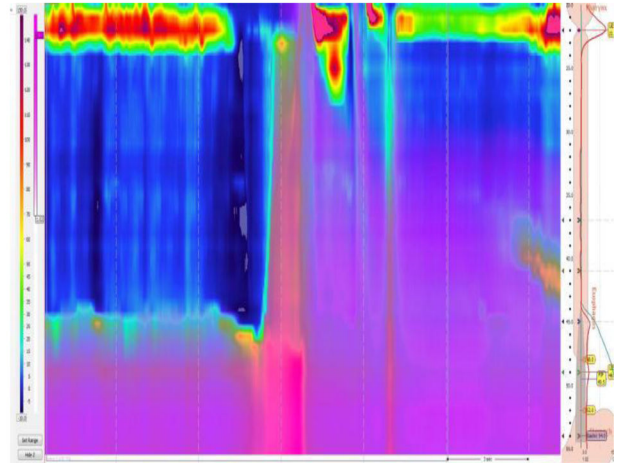


Figure 2. Differences in Rumination Identified on MII-pH Compared to PPHRIM.

(a) A 50% decrease in baseline impedance with or without a drop in pH on MII-pH could lead to rumination mistakenly identified as a reflux episode and result in the misdiagnosis of GERD. (b) PPHRIM, however, can identify rumination using both impedance and pressure characteristics unique to rumination and not GERD: gastric pressure increased > 30 mmHg with a retrograde pressure gradient and flow of gastric refluxate identified by a 50% decrease in impedance.

Table 1.

Options for Diagnostic Impression for MII-pH and PPRHIM.

Presence of Behavioral Condition	Diagnostic Group
No Behavioral Condition	<ul style="list-style-type: none">• Normal• Abnormal reflux burden/GERD
Behavioral Condition	<ul style="list-style-type: none">• SGB and/or Rumination• SGB and/or Rumination + GERD

MII-pH, multichannel intraluminal impedance pH; PPRHIM, post-prandial high resolution impedance manometry; GERD, gastroesophageal reflux disease; SGB, supragastric belching

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Inter-Rater Agreement for MII-pH & PPHRIM for Diagnostic Group and Behavioral Condition.

	PPHRIM		MII-pH		Post hoc MII-pH	
	Kappa	Agreement	Kappa	Agreement	Kappa	Agreement
Diagnostic Group	0.03 (-0.12, 0.18)	Slight	0.32 (0.17, 0.47)	Fair	0.47 (0.30, 0.64)	Moderate
Behavioral Condition	-0.22 (-0.46, 0.02)	Poor	0.13 (-0.11, 0.37)	Slight	0.19 (-0.07, 0.45)	Slight

MII-pH, multichannel intraluminal impedance pH; PPHRIM, post-prandial high resolution esophageal manometry

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Diagnostic Agreement between MII-pH and PPHRIM from Rater Diagnostic Impressions

MII-pH	PPHRIM				Total
	Normal	GERD	GERD + Behavioral	Behavioral only	
Normal	6	0	11	3	20
GERD	1	2	3	2	8
GERD + Behavioral	1	0	9	3	13
Behavioral only	7	3	6	10	25
Total	15	5	28	18	66

Gray cells highlight impressions with concordance.

MII-pH, multichannel intraluminal impedance pH; PPHRIM, post-prandial high resolution esophageal manometry; GERD, gastroesophageal reflux disease