Yadlapati Rena (Orcid ID: 0000-0002-7872-2033)
Gyawali C. Prakash (Orcid ID: 0000-0002-3388-0660)
Chen Joan Weichun (Orcid ID: 0000-0001-9425-3209)
Katzka David A. (Orcid ID: 0000-0003-2307-2317)
Richter Joel E. (Orcid ID: 0000-0001-5841-8491)
Yadlapati Rena (Orcid ID: 0000-0002-7872-2033)

CLINICAL ROLE OF AMBULATORY REFLUX MONITORING IN PPI NON-RESPONDERS:

Recommendation Statements

Rena Yadlapati[1], Andrew J Gawron[2,3], C. Prakash Gyawali[4], Joan Chen[5], John Clarke[6], Ronnie Fass[7], Anand Jain[8], Kristle Lynch[9], Abraham Khan[10], Philip O. Katz[11], David A. Katzka[12, 13], Joel Richter[14], Felice Schnoll-Sussman[11], Stuart J. Spechler[15], Michael F. Vaezi[16], Marcelo Vela[17], John E. Pandolfino[18]

1. Division of Gastroenterology & Hepatology, University of California San Diego, La Jolla, CA, USA
2. University of Utah, Salt Lake City UT, USA
3. Salt Lake City VA Medical Center , Salt Lake City UT, USA
4. Division of Gastroenterology, Washington University School of Medicine, St. Louis, MO, USA
5. Division of Gastroenterology, University of Michigan, Ann Arbor MI, USA
6. Division of Gastroenterology, Stanford University, Palo Alto CA, USA
7. Division of Gastroenterology, MetroHealth System, Cleveland OH, USA
8. Division of Gastroenterology, Emory University, Atlanta GA, USA
9. Division of Gastroenterology & Hepatology, Perelman School of Medicine, University of Pennsylvania, Philadelphia PA, USA
10. Center for Esophageal Health, Division of Gastroenterology & Hepatology, NYU Langone Health, New York, NY, USA
11. Division of Gastroenterology, Weill-Cornell Medical Center, New York NY, USA
12. Division of Gastroenterology, Mayo Clinic, Rochester MN, USA
13. Division of Gastroenterology, Columbia Presbyterian Medical Center, New York NY, USA
14. Division of Gastroenterology, University of South Florida, Tampa FL, USA
15. Division of Gastroenterology, Baylor Scott & White Health, Dallas TX, USA
16. Division of Gastroenterology, Vanderbilt Medical Center, Nashville TN, USA
17. Division of Gastroenterology, Mayo Clinic, Phoenix AZ, USA
18. Division of Gastroenterology, Northwestern Feinberg School of Medicine, Chicago, IL, USA

Corresponding Author:
Rena Yadlapati MD MS
9500 Gilman Drive MC0956
La Jolla, CA 92093
Phone: 714-420-5115
Email: ryadlapati@health.ucsd.edu

Running Head: Ambulatory Reflux Monitoring in PPI-Non Responders

Manuscript Word Count: 4308

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/apt.17180

Acknowledgements: We would like to sincerely acknowledge Dr. Donald O. Castell, a pioneer in the field of GERD with numerous contributions to reflux testing that laid the foundations for the recommendations put forth in this document. Dr. Castell accepted the invitation to participate in this study prior to his passing.

Disclosures:
RY: Consultant: Medtronic (Institutional), Ironwood Pharmaceuticals (Institutional), Phathom Pharmaceuticals, Med DataLink; Research support: Ironwood Pharmaceuticals; Advisory Board with Stock Options: RJS Mediagnostix
CPG: Consultant: Medtronic, Diversatek, Ironwood, Iso-Thrive, Quintiles, Takeda, Johnson&Johnson
JC: Consulting agreement - Phathom Pharmaceuticals
John Clarke: Consultant: Alnylam, Isothrive, Medtronic, Phathom Pharmaceuticals, Pfizer, Regeneron, Sanofi
RF: Advisor - Takeda, Medtronic, Phathom Pharmaceuticals, Neurogastrx, GERDCar, Celexio, Evoke Pharma Speaker - Astrazeneca, Takeda, GI Supply, Eisai, Johnson & Johnson
KL: Consultant Medtronic, Takeda, LUCID
AK: Consultant: Medtronic
POK: Consultant: Phathom Pharma, Takeda
DAK: Consulting for Takeda, Celgene and Regeneron
JR: Consultant for Medtronic
FSS: Consultant: Ethicon, Interpace, Medtronic
SJS: Consultant for Phathom Pharmaceuticals, Takeda Pharmaceuticals, Ironwood Pharmaceuticals, Interpace Diagnostics, Castle Biosciences, and ISOThrive
MFV: Ironwood, Phathom, ISOthrive, Sanofi, Bayer, Medtronic, Diversatek. Legal-consultation in litigations relating to acid suppressive therapy; Patent on Mucosal Integrity testing
MV: Medtronic - Consultant, Diversatek - Research Support
JEP: Consultant: Medtronic, Diversatek, Ethicon/Torax, Endogastric solutions, Takeda, Astra Zeneca, Phathom, Neurogastrx, Ironwood
AJG, AJ: None:

Author Contribution: Study Design (RY, AJG, JEP); Study Oversight (RY, JEP); RAND moderator (AJG); RAND participant (all authors); Data Collection (RY, JEP); Interpretation & Analysis of Data (RY, JEP); Drafting of initial manuscript (RY, AJG, JEP); Revising article (all authors); Approval of final version (all authors)

Funding Support: RY, CPG, JEP are supported by NIH R01 DK092217-04 (PI: Pandolfino); RY is supported by NIH K23 DK125266 (PI: Yadlapati).

Abbreviations: RAND Appropriateness Method (RAM); proton pump inhibitor (PPI); gastro-esophageal reflux disease (GERD); Acid exposure time (AET)

Writing Assistance: None

Keywords: refractory GERD; high-resolution manometry; laryngopharyngeal reflux

SUMMARY/ABSTRACT:

Background: Optimal ambulatory reflux monitoring methodology in symptomatic reflux patients continues to be debated.

Aims: This US initiative utilized published literature, and expert opinion to develop recommendation statements addressing use of ambulatory reflux monitoring in clinical practice.

Methods: The RAND Appropriateness Method (RAM) was utilized among 17 experts with discussion, revision and two rounds of ranking of recommendation statements. Ambulatory reflux monitoring protocol, methodology and thresholds ranked as appropriate by ≥80% panelists met criteria for appropriateness.

Results: Prolonged (96-hour recommended) wireless pH-monitoring off proton pump inhibitor (PPI) was identified as the appropriate diagnostic tool to assess need for acid suppression in patients with unproven gastro-esophageal reflux disease (GERD) and persisting typical reflux symptoms despite single dose PPI. Acid exposure time (AET) <4.0% on all days of monitoring with negative reflux-symptom association excludes GERD and does not support ongoing PPI treatment. Conversely, AET>6.0% across ≥2 days is conclusive evidence for GERD and supports treatment for GERD, while AET>10% across ≥2 days identifies severe acid burden that supports escalation of anti-reflux treatment. In previously proven GERD, impedance-pH monitoring on PPI is helpful in defining refractory GERD and mechanisms of continued symptoms; the presence of <40 reflux events, AET<2.0% and a negative reflux-symptom association does not support escalation of anti-reflux treatment. In contrast, AET>4.0% and positive reflux-symptom association supports escalation of anti-reflux treatment, including use of invasive therapeutics.

Conclusions: Statements meeting appropriateness for average clinical care have been identified when utilizing reflux monitoring in patients with typical reflux symptoms and PPI non-response.

BACKGROUND

Ambulatory reflux monitoring performed off acid suppression is the current gold standard to objectively evaluate esophageal acid burden for the diagnosis of gastro-esophageal reflux disease (GERD) in patients with reflux symptoms without erosive features of reflux disease on upper gastrointestinal endoscopy. [1-4] However, uncertainty exists in choosing the optimal reflux monitoring test and interpretation parameters between two distinct available systems. One is the wireless pH capsule placed transorally with capability to measure distal esophageal pH for up to 96 hours. The second is a catheter based system placed transnasally which measures esophageal pH for up to 24 hours, with added capability to assess liquid and gas transit bidirectionally throughout the esophagus as well as baseline impedance if using impedance-pH catheter. [1-3] Diagnostic thresholds that stratify acid exposure time as normal versus abnormal remain unclear. Consequently, ambulatory reflux monitoring is performed and variably interpreted across providers and centers.

In recent years, well-designed randomized trials examining clinically meaningful outcomes have sought to address these clinical knowledge gaps. [5-7] To date, recommendations regarding choice of ambulatory reflux monitoring systems and interpretation methodology which incorporate recent high quality data are limited. Thus, the objective of this initiative was to utilize the formal validated RAND/UCLA Appropriateness Method among a group of experts to identify appropriate recommendations for the clinical role, choice, and interpretation of ambulatory reflux monitoring.

METHODS

Study Design

In this prospective study we employed the RAND/University of California, Los Angeles Appropriateness Method (RAM) to assess the appropriateness of metrics related to ambulatory reflux monitoring. This study was supported by an overarching grant aimed to determine the effectiveness of physiologic testing in patients with GERD symptoms who are not responsive to proton pump inhibitor (PPI) therapy [NIH R01 DK092217-04].

RAND/University of California, Los Angeles Appropriateness Method

When using RAM, the concept of appropriateness refers to the relative weight of the benefits and harms of an intervention. An appropriate statement is one where the expected health benefit exceeds the expected negative consequences by a sufficiently wide margin exclusive of costs.[8] RAM utilizes a modified Delphi method that, unlike the original Delphi, provides panelists with the opportunity to discuss their judgments and responses between rating rounds and at a face-to-face meeting, similar to the National Institute of Health Consensus Conferences. This methodology is also applicable when randomized controlled trials are not available or cannot provide evidence at a level of detail sufficient to apply to the wide range of patients seen in everyday clinical practice.[8] This is a well-described method used to develop quality indicators which has applicability across a broad range of disease processes and procedures, and across multiple countries.[9-19]

Recruitment of the Expert Panel

At the time of grant proposal submission in 2016 [NIH R01 DK092217-04], adult gastroenterologists with a clinical focus on the evaluation and management of GERD were invited to participate as expert panelists for this planned study. The main selection criteria in the

nomination process included leadership within the field of GERD, and diversity in age, gender and geographical location in the US. While the ideal sample size for members in such expert panels has not been defined, RAM experts suggest that the panels can be of any size that permits sufficient diversity (a minimum of 7), while ensuring that all members have a chance to participate.[8] An electronic invitation to participate in this study was sent to the panel of experts briefly describing the study objectives, description of the RAM, and responsibilities of each expert panel member.

All invited panelists accepted and participated in the process, led by two co-chairs (RY & JEP), and a moderator with a health services research background and experience with RAM (AJG). The 17 member panel was comprised of 4 females and 13 males with a mean of 19.2 years in practice (SD 12.9). On average the panel cared for 53.9 patients (SD 92.7) a month with typical GERD symptoms and 24.9 patients (SD 18.6) a month with atypical GERD symptoms.

Round 1: Initial ranking of recommendation statements

The co-chairs and moderator proposed recommendation statements within the following domains: role and protocol of wireless reflux monitoring, thresholds for interpretation of prolonged wireless reflux monitoring off PPI therapy, and thresholds for interpretation of impedance-pH monitoring on PPI therapy. Acid exposure time (AET) was the primary threshold assessed as the measure of esophageal acid burden given that prior data identifies that AET performs comparably though better than other metrics such as DeMeester score, dominant pattern or acid exposure trajectory. All panelists were sent a document detailing the objectives, supportive literature, instructions and a link to a REDCap survey instrument via email. The instructions highlighted that the recommendation statements did not necessarily have to apply to any one specific patient, but rather, they were relevant to the overall care of patients

undergoing ambulatory reflux monitoring. A recommendation was considered appropriate if adherence to this recommendation was critical to quality ambulatory reflux monitoring exclusive of cost or feasibility, with applicability to the average patient presenting to the average physician at an average practice setting. We emphasized that the panel members should not consider cost implications or the feasibility of implementing the recommendation in their rankings. Each recommendation statement was ranked on a 9-point interval scale in which a score of 1-3 was signified as inappropriate, 4-6 was of uncertain appropriateness and 7-9 was deemed appropriate. The panelists were provided the opportunity to include comments regarding each proposed recommendation and to suggest modifications.

Round 2 meeting: Discussion of potential quality indicators, re-wording, and re-ranking

The Round 2 face-to-face meeting among the expert panel members was conducted virtually in September 2021 and led by the moderator (AJG). For each proposed recommendation statement, the panelists reviewed the aggregated ranking results from Round 1, discussed available evidence and expert opinions/experiences, proposed re-wording of recommendations and suggested new recommendations. Following the meeting, panel members independently re-ranked each of the proposed recommendations for their perceived level of appropriateness.

Outcome & Analysis

The primary outcome was appropriateness of each intervention based on the recommendation statements. Final appropriateness was based on median rankings and the dispersion of rankings. Per RAND constructs, agreement required 80% or more of panelists' rankings in the same three point range: Inappropriate (1–3), Equivocal/Uncertain (4–6), or Appropriate (7–9). Disagreement was present when more than 20% of the rankings were in disparate categories. The moderators emphasized to the panelists that the objective of this

process was not to necessarily achieve consensus, but rather to highlight areas of agreement, inconsistency, and disagreement.

RESULTS

Eleven recommendation statements were proposed in Round 1. During Round 2 all recommendations underwent re-wording by the panel group and 4 new recommendations were added by the panel group for a total of fifteen proposed recommendation statements. Ranking following round 2 resulted in a final 8 recommendation statements meeting criteria as appropriate, and the remaining 7 as equivocal/indeterminate. None of the recommendations were ranked as inappropriate.

Clinical Role & Protocol for Wireless pH Monitoring

- **Prolonged wireless pH monitoring off PPI is the preferred diagnostic tool to assess need for acid suppression in patients with unproven GERD and typical reflux symptoms (heartburn/regurgitation) not adequately responding to single dose PPI therapy. [100% Ranked as Appropriate; Median Score 8]**
- **The preferred duration for wireless pH monitoring off acid suppression is 96 hours. [88% Ranked as Appropriate; Median Score 8]**

Up to 50% of patients with typical esophageal symptoms of reflux (heartburn and/or regurgitation) will not derive adequate symptom relief with an empiric trial of PPI therapy. When PPI non-response is encountered, upper GI endoscopy off PPI is recommended to evaluate for mucosal evidence of reflux disease such as Los Angeles grade B, C or D erosive esophagitis, Barrett's esophagus, or peptic stricture, or non-GERD related esophageal disorders, such as eosinophilic esophagitis, lymphocytic esophagitis, and others.[3] However, up to 80% of symptomatic patients will have normal healthy appearing esophageal mucosa. In this common scenario of visually normal esophageal mucosa, ambulatory reflux monitoring off acid suppression is recommended to quantify esophageal acid burden. Wireless pH capsule monitoring performed off therapy is the preferred system for evaluation of esophageal acid burden in PPI non-responder patients with typical reflux symptoms, in contrast to transnasal

catheter based monitoring (regardless of whether pH only or impedance-pH monitoring is considered). The preference for wireless pH monitoring relates to the ability of this technique to monitor pH over durations greater than 24 hours, with improved patient tolerance, and greater diagnostic accuracy. [6, 7, 20-28]

The optimal duration of prolonged wireless pH monitoring has been a frequent topic of debate. Based on the battery life of the data receiver device accompanying the wireless pH capsule, monitoring for up to 96 hours is possible. However, varying durations of monitoring are employed in the clinical setting across community and academic settings, ranging from 48 to 96 hours. Thus, the recommended duration of monitoring was extensively discussed during Round 2. Emerging published data highlights the significant variability in acid exposure from one day to another, as well as higher levels of acid exposure frequently seen on day 1 of monitoring compared to other days. [7, 20, 22-24, 27, 29-34] Specifically, recent data from the double-blinded clinical trial under this overarching grant [NIH R01 DK092217-04] were reviewed which identified prognostic performance of wireless reflux monitoring was significantly lower when data from the first 48 hours was assessed alone compared to 96 hours (Area under curve 48 hours 0.57 vs 96 hours 0.63; p=0.01) and also noted a significantly higher AET on day 1 of monitoring compared to the other days. [6, 7] As a result, 48 hours of pH monitoring risks a high false positive rate as well as the unclear significance of studies with discordant results from day 1 to day 2. Overall these studies highlight that 96-hour monitoring is more optimal than shorter durations of monitoring in predicting discontinuation versus ongoing need for PPI therapy. Therefore, the panel advocated for 96 hours of monitoring (88% appropriate) whereas 48 hours of monitoring was ranked as indeterminate by 35%, inappropriate by 47%, and appropriate by 18% of the panel.

Diagnostic Thresholds for Prolonged Wireless pH Monitoring off PPI Therapy

Acid Exposure Time Thresholds

- **In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain not responsive to single dose PPI, an acid exposure time less than 4.0% on all days of monitoring and an overall negative reflux-symptom association on prolonged wireless pH monitoring off PPI therapy does not support treatment of GERD with a PPI. [94% Ranked as Appropriate; Median Score 9]**

Acid exposure refers to measured esophageal intraluminal pH of less than 4.0. AET is the commonly utilized metric for esophageal acid exposure, calculated as the percentage of time the pH is less than 4.0. Although the Lyon Consensus recommends that an AET less than 4.0% is considered physiologic, a multitude of AET thresholds are used in clinical practice to define a "normal" or "physiologic" study, [2] ranging anywhere from 4.0% to 6.0%. A recent study examining multiple metrics on wireless pH monitoring including AETs of 4.0%, 5.0% and 6.0% identified that an AET of <4.0% has the greatest predictive value for a patient's ability to discontinue PPI therapy while maintaining a minimal symptom burden. A significantly greater proportion of patients with an overall AET < 4.0% were able to discontinue PPI therapy compared to those with an overall AET > 4.0%. Further, the number of days with an AET < 4.0% was of prognostic value, where the odds of PPI discontinuation were 10 times greater when AET was less than 4.0% across all days of monitoring. [6] Therefore, the panel agreed that in patients without previously established GERD who report heartburn, regurgitation and/or non-cardiac chest pain despite single dose PPI, an AET of less than 4.0% across all days of prolonged wireless pH monitoring off PPI therapy is consistent with a normal study for whom treatment of GERD with a PPI is not recommended.

- **In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain not responsive to single dose PPI, an acid exposure time greater than 6.0% across 2 or more days is diagnostic of GERD and supports treatment for GERD. [100% Ranked as Appropriate; Median Score 9]**

The threshold to define pathologic acid exposure was discussed extensively. Consistent with the Lyon Consensus, the group agreed that an AET between 4.0 to 6.0% likely represents a borderline range of GERD. [2] Only 71% of the panel ranked AET>4.0% across 2 or more days as diagnostic of GERD, with a median score of 6, which did not meet RAM criteria for appropriateness (Supplemental Table); the discussion in round 2 defaulted to the statement above as being appropriate. Thus, when AET is between 4.0 and 6.0%, further clinical consideration and additional clinical data are preferred to determine the need for GERD management, since other factors such as reflux hypersensitivity, motility disorders and behavioral disorders such as supragastric belching and rumination may be contributing to patient symptoms.

Consistent with Lyon Consensus, the panel agreed that an AET greater than 6.0% is reflective of pathologic esophageal acid burden. [2] Acknowledging day to day variability, the expert group unanimously agreed that an AET of 6.0% or greater on at least two days of pH monitoring is diagnostic of GERD supporting the use of standard GERD therapies including acid suppressive agents.

- **In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain and proven GERD, an acid exposure time greater than 10% across 2 or more days on prolonged wireless pH monitoring off acid suppression is consistent with severe acid burden and supports escalation of anti-reflux treatment. [94% Agreement; Median Score 9]**

The panel also recognized the importance of identifying patients with severe esophageal acid burden that may not derive adequate symptom relief with standard GERD therapies. Data from the recent study by Yadlapati and colleagues were reviewed which identified an AET of 10.3% as the lowest AET which maintained at least 90% specificity in predicting PPI discontinuation. [7] Therefore, the panel agreed that an AET greater than 10% across two or more days reflects

severe acid burden that may require escalation of medical anti-reflux treatment or utilization of endoscopic or surgical interventions. [35]

Reflux Symptom Association

In addition to acid exposure, ambulatory reflux monitoring also reports the relationship between patient symptoms and reflux episodes. Metrics of reflux symptom association include the symptom association probability and symptom index. Generally, positive symptom association probability (>95%) and symptom index greater than 50% increases confidence that a patient's symptoms are related to gastroesophageal reflux. [2] However, negative reflux symptom association is less convincing as reflux symptom association measurement relies on prompt patient report of perceived symptoms within a 2-minute window. Rome IV posits that a positive reflux symptom association in the absence of elevated esophageal acid exposure signifies reflux hypersensitivity. [36] However some members of this current panel did not feel compelled to distinguish between a negative study or reflux hypersensitivity on the basis of reflux symptom association. Further the results from the overarching grant were reviewed which failed to identify a significant association between symptom index or symptom association probability and ability to discontinue PPI. [7] A statement proposing a diagnosis of reflux hypersensitivity when positive reflux symptom association is encountered in conjunction with acid exposure time less than 4.0% on all days of monitoring did not meet agreement as appropriate (Supplemental Table). Similarly, the group disagreed regarding the relevance of a positive reflux symptom association in the setting of elevated acid exposure (AET > 6.0% across 2 or more days). Some panelists felt that a positive reflux symptom association in this setting signifies a higher likelihood of symptom response to treatment whereas other panelists felt that there was insufficient data to merit this recommendation (Supplemental Table).

Impedance-pH Monitoring on PPI Therapy

- **In patients with proven GERD, 24 hour pH impedance on PPI therapy is helpful in defining refractory GERD and mechanisms of continued symptoms. (88% Ranked as Appropriate; Median Score: 8)**
- **In patients with proven GERD with ongoing symptoms despite optimized PPI therapy undergoing pH impedance monitoring on double dose PPI therapy, the presence of fewer than 40 reflux events, an acid exposure time less than 2.0% and a negative reflux-symptom association does not support escalation of anti-reflux treatment. (100% Ranked as Appropriate; Median Score: 9).**
- **In patients with proven GERD with ongoing symptoms despite optimized PPI therapy undergoing pH impedance monitoring on double dose PPI therapy, the presence of an acid exposure time greater than 4.0% and positive reflux-symptom association supports escalation of anti-reflux treatment. (94% Ranked as Appropriate; Median Score: 8).**

For patients with already proven GERD (prior erosive reflux disease or positive ambulatory reflux monitoring study performed off PPI) and non-response to optimized PPI therapy (double dose before-meal PPI) impedance-pH monitoring performed on PPI therapy, while not mandatory, is an important test to assess for underlying mechanisms for GERD refractoriness.[37] Few studies have examined impedance-pH metrics on PPI therapy that correlate with treatment outcomes in patients with GERD. A recent study highlights that 40 may be a relevant threshold for number of reflux episodes on impedance-pH monitoring, and that an AET > 4.0% while on PPI therapy may be associated with patients more likely to respond to surgical or endoscopic anti-reflux intervention.[37] Therefore on the basis of limited data and expert experiences, the panel agreed that an AET less than 2.0%, fewer than 40 reflux events and a negative reflux symptom association on 24 hour impedance-pH monitoring performed on PPI therapy does not support an escalation of anti-reflux treatment.[38] For these patients, ongoing

symptoms may be related to other non-GERD factors and escalation of GERD therapy is not justified. On the other hand, refractory GERD may manifest as an AET greater than 4.0% and a positive reflux-symptom association on 24-hour impedance-pH monitoring performed on PPI therapy. For instance, a landmark randomized trial demonstrated that patients with heartburn and positive reflux symptom association on pH impedance on PPI therapy had better response to anti-reflux surgery compared to medical treatment [5]. Therefore, in patients with significantly refractory GERD it is reasonable to escalate GERD management. [5]

The various permutations of AET between 2.0 to 4.0% and reflux burden between 40 to 80 reflux events were also discussed, and while many ranked high, none of the statements pertaining to these thresholds met agreement as appropriate recommendations (Supplemental Table).

DISCUSSION

This US initiative utilized RAND appropriateness methodology to develop recommendation statements regarding the clinical utilization and interpretation of ambulatory reflux monitoring in patients with typical reflux symptoms. The recommendation statements drafted for this study were intended to address two unmet needs within the field of GERD diagnosis and management: to provide guidance to clinicians on the protocol and interpretation of reflux monitoring, and to develop standardized criteria for non-erosive GERD for patient management and guide future study designs to minimize heterogeneity of study populations. By combining RAM with expert opinion from a diverse nationwide representative cohort of GERD experts, the study concluded with excellent agreement among the expert panel that wireless pH monitoring performed off PPI over 96 hours represents the most appropriate protocol for investigation of typical reflux symptoms persisting despite standard PPI therapy. Additionally, the experts overwhelmingly agreed that impedance-pH monitoring performed on PPI can identify PPI refractory GERD appropriate for escalation of management in symptomatic patients with previously proven GERD. Finally, the most appropriate diagnostic thresholds for diagnosis of GERD were identified, both for off-PPI and on-PPI studies, while thresholds and metrics that remain inconclusive were defined.

In terms of interpretation, the experts agreed that a wireless pH monitoring study off PPI with an acid exposure time less than 4.0% across all days of monitoring indicates a very low likelihood of GERD, where acid suppression is not recommended. On the other hand, a study with acid exposure time greater than 6.0% on two or more days of pH monitoring indicates a high likelihood of GERD. Furthermore, higher levels of acid exposure (10% or greater) suggest more severe GERD and a high likelihood that GERD management may need to be escalated. Although a majority (71%) of panelists agreed that an acid exposure time greater than 4.0% on 2 or more days of prolonged reflux monitoring was diagnostic of GERD, the ranking did not

meet criteria for appropriateness. Some panelists expressed that an acid exposure time between 4.0 to 6.0% should still be considered inconclusive, as per the Lyon Consensus. The refinement of this threshold (acid exposure time 4.0% to 6.0%, as well as number of days with an acid exposure time greater than 4.0%) is a priority for future studies.

The experts agreed that patients with proven GERD and ongoing symptoms despite optimized PPI therapy could benefit from impedance-pH monitoring on PPI since findings could demonstrate if the ongoing symptoms are related to reflux. In terms of interpretation, a pH impedance study on PPI with less than 40 reflux events, acid exposure time less than 2.0%, and a negative reflux symptom association indicates PPI controlled GERD and potential for alternative non-GERD etiology of ongoing symptoms. On the other hand, a pH-impedance study on PPI with an acid exposure time greater than 4.0% and a positive reflux symptom association indicates PPI refractory GERD and supports escalation of GERD therapy.

According to international recommendations, wireless pH or pH-impedance monitoring off PPI can be used to assess for non-erosive GERD in patients with typical reflux symptoms. The flexibility in choice of reflux monitoring relates to lack of availability of wireless pH monitoring in some countries outside of the US. [4] While recent guidelines endorse prolonged reflux monitoring, they however fail to define the recommended duration of monitoring. Consequently, varying durations are utilized from 48 to 96 hours, even though results from a 48-hour study have been shown to be discordant from a 96-hour study. [23 7] Given availability and payor coverage of both systems in the US, the panelists weighed the advantages and disadvantages of both systems, and determined that prolonged wireless pH monitoring off PPI therapy is the preferred diagnostic tool for non-erosive reflux disease in patients with typical symptoms such as heartburn and regurgitation not responsive to PPI therapy. Thus, standardizing the duration of prolonged reflux monitoring was a priority and a heavily debated topic. In the end the superior diagnostic yield of data from 96 hours of monitoring was felt to

outweigh potential constraints on resource availability. pH-impedance off PPI therapy is an alternative for patients who cannot undergo wireless pH monitoring (e.g., pacemaker, nickel allergy, inability to undergo a sedated upper GI endoscopy) or in scenarios where wireless pH monitoring is not accessible. Further, while the clinical evaluation of patients with atypical symptoms was beyond the scope of this initiative, it is important to note that recent guidelines suggest a role of up front pH-impedance monitoring for the evaluation of extra-esophageal GERD and that pH-impedance monitoring is of particular value in instances where weakly acidic reflux is relevant such as extra-esophageal reflux and belching disorders [1, 3].

The panelists generally agreed with the acid exposure time thresholds defined by the Lyon Consensus, and additionally provided recommendations on diagnostic interpretation of acid exposure across multiple days of monitoring [2]. Importantly, a physiologic acid exposure (AET<4.0%) across every day of monitoring was considered consistent with a normal study with the implication that patients with physiologic AET should be titrated off PPI therapy. In these cases, esophageal symptoms may be related to a functional esophageal disorder, a behavioral disorder, esophageal motility disorder or other process. On the other hand, patients with pathologic acid exposure (AET>6.0%) on two or more days of monitoring are expected to benefit from optimized lifestyle and pharmacologic anti-reflux management. Patients with very high levels of acid exposure (AET>10%) may be less responsive to only lifestyle and pharmacologic management, particularly in the setting of a large hiatal hernia and/or bipositional/nocturnal GERD, and may require escalation of GERD management. [39]

There was less certainty regarding standardization of the interpretation of pH-impedance on PPI therapy. Based on recent data, the group agreed on two recommendations regarding pH impedance on PPI, the definition of PPI controlled GERD (AET < 2.0%, <40 reflux events, and a negative reflux symptom association) and a definition of PPI refractory GERD (AET > 4.0% and a positive reflux symptom association). [5, 37] Further, this initiative focused on well-established

and clinically utilized metrics of reflux monitoring such as acid exposure time and number of reflux events. Since the initiative focus was to provide recommendations to the general clinician interpreting reflux monitoring studies the utility of novel impedance-pH metrics such as post-reflux swallow peristaltic wave index, which are not automated or easily interpretable, were not discussed. Nonetheless, as per the Lyon consensus and other recent guidelines, the panel agrees that mean nocturnal baseline impedance and post-reflux swallow peristaltic wave index are of value, particularly for the inconclusive GERD scenarios.[40 2] While beyond the scope of this initiative, it is important to note the critical importance of esophageal physiologic tests such as high resolution manometry to exclude achalasia in patients undergoing evaluation for anti-reflux surgery.[1, 4, 41]

This initiative highlighted areas in need of further investigation and clarity. Although reflux symptom association parameters are commonly used to assess for reflux hypersensitivity, the group could not agree on the clear clinical relevance of reflux symptom association. Future research to understand the distinctions in treatment outcome between patients with functional esophageal disorders and reflux hypersensitivity are needed. The stark contrast in confidence regarding thresholds on wireless pH monitoring off PPI compared to thresholds on pH-impedance monitoring on PPI highlights the need to better understand the clinical role of pH-impedance monitoring on PPI.

A major strength of this study consists of wide and diverse representation of esophageal experts from different practice settings, who participated in development of potential statements, debate and review of published data and determining appropriateness of the recommended statements. The RAND process has been widely used in prior studies and allows for a rigorous approach to revealing areas of agreement and disagreement in clinical care and to present knowledge gaps for future generation of evidence.[42, 43] Assessed domains were focused toward existing clinical dilemmas. Given the focused scope of this study, we did not assess statements

relating to extraesophageal symptoms, the role of esophageal physiologic testing beyond reflux monitoring, or emerging metrics from reflux monitoring that are not yet widely used (e.g., post-swallow peristaltic wave index or mean nocturnal baseline impedance). Based on data from our prior study highlighting that total acid exposure time is the best performing physio-marker, we did not assess whether the occurrence of increased acid exposure should be viewed differently depending on its presence in the upright and/or supine position. We also did not assess composite metrics such as the DeMeester score, dominant pattern of acid exposure or acid exposure trajectory; prior data supports that acid exposure time has comparable yet still better performance in predicting ability to discontinue PPI therapy compared to the aforementioned composite metrics.[6] Actual patient scenarios were not utilized, and the experts were required to generalize the statements across average patients seen by an average gastroenterologist, which may have influenced some of the responses. Nevertheless, we feel the process has led to a better understanding of how GI experts view the current landscape of GERD diagnostics in the US, and which test protocol and metrics are best suited for evaluation of the symptomatic GERD patient.

In summary, a diverse US group of GI expert GERD panelists concluded that a 96-hour wireless pH monitoring study off PPI is most appropriate for further evaluation of typical GERD symptoms not responding to PPI, while findings from impedance-pH monitoring on PPI could identify PPI-refractory GERD in symptomatic patients with previously proven GERD. These recommendations provide a framework for approaching reflux monitoring in patients with typical reflux symptoms and PPI non-response.

Table 1. Final Recommendation Statements		
FINAL RECOMMENDATION STATEMENTS	Median (Range)	% Agreement
Role/Protocol of Wireless pH Monitoring		
Prolonged wireless pH monitoring off PPI is the preferred diagnostic tool to assess need for acid suppression in patients with unproven GERD and typical reflux symptoms (heartburn/regurgitation) not adequately responding to single dose PPI therapy.	8 (7, 9)	100% Appropriate
The preferred duration for wireless pH monitoring off acid suppression is 96 hours.	8 (4, 9)	88% Appropriate
Prolonged Wireless pH Monitoring off PPI - Thresholds		
In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain not responsive to single dose PPI, an acid exposure time less than 4.0% on all days of monitoring and an overall negative reflux-symptom association on prolonged wireless pH monitoring off PPI therapy does not support treatment of GERD with a PPI.	9 (2, 9)	94% Appropriate
In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain not responsive to single dose PPI, an acid exposure time greater than 6.0% across 2 or more days is diagnostic of and supports treatment for GERD.	9 (7, 9)	100% Appropriate
In patients with symptoms of heartburn, regurgitation, and/or non-cardiac chest pain and proven GERD, an acid exposure time greater than 10% across 2 or more days on prolonged wireless pH monitoring off acid suppression is consistent with severe acid burden and supports escalation of anti-reflux treatment.	9 (6, 9)	94% Appropriate
pH impedance on PPI - Thresholds		
In patients with proven GERD, 24 hour pH impedance on PPI therapy is helpful in defining refractory GERD and mechanisms of continued symptoms	8 (4, 9)	88% Appropriate
In patients with proven GERD with ongoing symptoms despite optimized PPI therapy undergoing pH impedance monitoring on double dose PPI therapy, the presence of fewer than 40 reflux events, an acid exposure time less than 2.0% and a negative reflux-symptom association does not support escalation of anti-reflux treatment.	9 (7, 9)	100% Appropriate
In patients with proven GERD with ongoing symptoms despite optimized PPI therapy undergoing pH impedance monitoring on double dose PPI therapy, the presence of an acid exposure time greater than 4.0% and positive reflux-symptom association supports escalation of anti-reflux treatment.	8 (5, 9)	94% Appropriate

FIGURE LEGEND

Figure 1. Ambulatory Reflux Monitoring Protocol & Interpretation Scheme to Assess Patients with Unproven GERD and Typical Reflux Symptoms Persistent Despite PPI Therapy. The optimal protocol is 96 hours reflux monitoring off acid suppression. GERD is excluded when acid exposure time (AET) is less than 4.0% on all days of monitoring, in which case ongoing acid suppression is not indicated. GERD is diagnosed when AET is greater than 6.0% on 2 or more days of monitoring. AET patterns in between represent an inconclusive diagnosis, in which case other clinical and diagnostic data can help strengthen confidence in a diagnostic impression. Finally, 24 hour pH-impedance may be of value for patients with proven GERD and ongoing symptoms despite management to evaluate for refractory GERD.

REFERENCES

1. Gyawali CP, Carlson DA, Chen JW, et al. ACG Clinical Guidelines: Clinical Use of Esophageal Physiologic Testing. Am J Gastroenterol 2020;115:1412-1428.
2. Gyawali CP, Kahrilas PJ, Savarino E, et al. Modern diagnosis of GERD: the Lyon Consensus. Gut 2018;67:1351-1362.
3. Katz PO, Dunbar KB, Schnoll-Sussman FH, et al. ACG Clinical Guideline for the Diagnosis and Management of Gastroesophageal Reflux Disease. Am J Gastroenterol 2021.
4. Zerbib F, Bredenoord AJ, Fass R, et al. ESNM/ANMS consensus paper: Diagnosis and management of refractory gastro-esophageal reflux disease. Neurogastroenterol Motil 2020:e14075.
5. Spechler SJ, Hunter JG, Jones KM, et al. Randomized Trial of Medical versus Surgical Treatment for Refractory Heartburn. N Engl J Med 2019;381:1513-1523.
6. Yadlapati R, Masihi M, Gyawali CP, et al. Ambulatory Reflux Monitoring Guides Proton Pump Inhibitor Discontinuation in Patients With Gastroesophageal Reflux Symptoms: A Clinical Trial. Gastroenterology 2021;160:174-182 e1.
7. Yadlapati R, Gyawali PC, Masihi M, et al. OPTIMAL WIRELESS REFLUX MONITORING METRICS TO PREDICT DISCONTINUATION OF PROTON PUMP INHIBITOR THERAPY. Am J Gastroenterol 2022.
8. Fitch K, Aguilar MD, Burnand B, et al. The RAND/UCLA appropriateness method user's manual. Santa Monica: RAND 2001.
9. Yadlapati R, Gawron AJ, Keswani RN, et al. Identification of Quality Measures for Performance of and Interpretation of Data From Esophageal Manometry. Clin Gastroenterol Hepatol 2016;14:526-534 e1.
10. Yadlapati R, Gawron AJ, Bilimoria K, et al. Development of quality measures for the care of patients with gastroesophageal reflux disease. Clin Gastroenterol Hepatol 2015;13:874-83 e2.
11. Shekelle PG, Kahan JP, Bernstein SJ, et al. The reproducibility of a method to identify the overuse and underuse of medical procedures. N Engl J Med 1998;338:1888-95.
12. Halverson AL, Sellers MM, Bilimoria KY, et al. Identification of process measures to reduce postoperative readmission. J Gastrointest Surg 2014;18:1407-15.
13. Bilimoria KY, Bentrem DJ, Lillemoe KD, et al. Assessment of pancreatic cancer care in the United States based on formally developed quality indicators. J Natl Cancer Inst 2009;101:848-59.
14. Bilimoria KY, Raval MV, Bentrem DJ, et al. National assessment of melanoma care using formally developed quality indicators. J Clin Oncol 2009;27:5445-51.
15. Aguilar MD, Fitch K, Lazaro P, et al. The appropriateness of use of percutaneous transluminal coronary angioplasty in Spain. Int J Cardiol 2001;78:213-21; discussion 221-3.
16. Bernstein SJ, Lazaro P, Fitch K, et al. Effect of specialty and nationality on panel judgments of the appropriateness of coronary revascularization: a pilot study. Med Care 2001;39:513-20.
17. Lawson EH, Gibbons MM, Ko CY, et al. The appropriateness method has acceptable reliability and validity for assessing overuse and underuse of surgical procedures. J Clin Epidemiol 2012;65:1133-43.
18. Maggard MA, McGory ML, Shekelle PG, et al. Quality indicators in bariatric surgery: improving quality of care. Surg Obes Relat Dis 2006;2:423-9; discussion 429-30.
19. McGory ML, Shekelle PG, Ko CY. Development of quality indicators for patients undergoing colorectal cancer surgery. J Natl Cancer Inst 2006;98:1623-33.
20. Pandolfino JE, Richter JE, Ours T, et al. Ambulatory esophageal pH monitoring using a wireless system. Am J Gastroenterol 2003;98:740-9.

21. Wiener GJ, Morgan TM, Copper JB, et al. Ambulatory 24-hour esophageal pH monitoring. Reproducibility and variability of pH parameters. Dig Dis Sci 1988;33:1127-33.
22. Ahlawat SK, Novak DJ, Williams DC, et al. Day-to-day variability in acid reflux patterns using the BRAVO pH monitoring system. J Clin Gastroenterol 2006;40:20-4.
23. Hasak S, Yadlapati R, Altayar O, et al. Prolonged Wireless pH Monitoring in Patients With Persistent Reflux Symptoms Despite Proton Pump Inhibitor Therapy. Clin Gastroenterol Hepatol 2020.
24. Yadlapati R, Ciolino JD, Craft J, et al. Trajectory assessment is useful when day-to-day esophageal acid exposure varies in prolonged wireless pH monitoring. Dis Esophagus 2019;32.
25. Penagini R, Sweis R, Mauro A, et al. Inconsistency in the Diagnosis of Functional Heartburn: Usefulness of Prolonged Wireless pH Monitoring in Patients With Proton Pump Inhibitor Refractory Gastroesophageal Reflux Disease. J Neurogastroenterol Motil 2015;21:265-72.
26. Scarpulla G, Camilleri S, Galante P, et al. The impact of prolonged pH measurements on the diagnosis of gastroesophageal reflux disease: 4-day wireless pH studies. Am J Gastroenterol 2007;102:2642-7.
27. Sweis R, Fox M, Anggiansah A, et al. Prolonged, wireless pH-studies have a high diagnostic yield in patients with reflux symptoms and negative 24-h catheter-based pH-studies. Neurogastroenterol Motil 2011;23:419-26.
28. Connor J, Richter J. Increasing yield also increases false positives and best serves to exclude GERD. Am J Gastroenterol 2006;101:460-3.
29. Hirano I, Zhang Q, Pandolfino JE, et al. Four-day Bravo pH capsule monitoring with and without proton pump inhibitor therapy. Clin Gastroenterol Hepatol 2005;3:1083-8.
30. Lacy BE, Dukowicz AC, Robertson DJ, et al. Clinical utility of the wireless pH capsule. J Clin Gastroenterol 2011;45:429-35.
31. Pandolfino JE, Kwiatek MA. Use and utility of the Bravo pH capsule. J Clin Gastroenterol 2008;42:571-8.
32. Prakash C, Clouse RE. Value of extended recording time with wireless pH monitoring in evaluating gastroesophageal reflux disease. Clin Gastroenterol Hepatol 2005;3:329-34.
33. Roman S, Mion F, Zerbib F, et al. Wireless pH capsule--yield in clinical practice. Endoscopy 2012;44:270-6.
34. Wenner J, Johnsson F, Johansson J, et al. Wireless oesophageal pH monitoring: feasibility, safety and normal values in healthy subjects. Scand J Gastroenterol 2005;40:768-74.
35. Krill JT, Naik RD, Higginbotham T, et al. Association Between Response to Acid-Suppression Therapy and Efficacy of Antireflux Surgery in Patients With Extraesophageal Reflux. Clin Gastroenterol Hepatol 2017;15:675-681.
36. Aziz Q, Fass R, Gyawali CP, et al. Functional Esophageal Disorders. Gastroenterology 2016.
37. Gyawali CP, Tutuian R, Zerbib F, et al. Value of pH Impedance Monitoring While on Twice-Daily Proton Pump Inhibitor Therapy to Identify Need for Escalation of Reflux Management. Gastroenterology 2021;161:1412-1422.
38. Patel A, Sayuk GS, Gyawali CP. Parameters on esophageal pH-impedance monitoring that predict outcomes of patients with gastroesophageal reflux disease. Clin Gastroenterol Hepatol 2015;13:884-91.
39. Rusu RI, Fox MR, Tucker E, et al. Validation of the Lyon classification for GORD diagnosis: acid exposure time assessed by prolonged wireless pH monitoring in healthy controls and patients with erosive oesophagitis. Gut 2021;70:2230-2237.
40. Visaggi P, Mariani L, Svizzero FB, et al. Clinical use of mean nocturnal baseline impedance and post-reflux swallow-induced peristaltic wave index for the diagnosis of gastro-esophageal reflux disease. Esophagus 2022.

41. Yadlapati R, Gyawali CP, Pandolfino JE, et al. AGA Clinical Practice Update on the Personalized Approach to the Evaluation and Management of GERD: Expert Review. Clin Gastroenterol Hepatol 2022;20:984-994 e1.
42. Gawron AJ, Bell R, Abu Dayyeh BK, et al. Surgical and endoscopic management options for patients with GERD based on proton pump inhibitor symptom response: recommendations from an expert U.S. panel. Gastrointest Endosc 2020;92:78-87 e2.
43. 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;113:980-986.

Ambulatory Reflux Monitoring Protocol & Interpretation Scheme to Assess Patients with Unproven GERD and Typical Reflux Symptoms Persistent Despite PPI Therapy