ACG Clinical Guidelines: Clinical Use of Esophageal Physiologic Testing

C. Prakash Gyawali, MD, MRCP, FACC\textsuperscript{1}, Dustin A. Carlson, MD\textsuperscript{2}, Joan W. Chen, MD\textsuperscript{3}, Amit Patel, MD\textsuperscript{4}, Robert J. Wong, MD, MS, FACC (GRADE Methodologist)\textsuperscript{5} and Rena H. Yadlapati, MD, MSHS\textsuperscript{6}

Esophageal symptoms are common and may indicate the presence of gastroesophageal reflux disease (GERD), structural processes, motor dysfunction, behavioral conditions, or functional disorders. Esophageal physiologic tests are often performed when initial endoscopic evaluation is unrevealing, especially when symptoms persist despite empiric management. Commonly used esophageal physiologic tests include esophageal manometry, ambulatory reflux monitoring, and barium esophagram. Functional luminal imaging probe (FLIP) has recently been approved for the evaluation of esophageal pressure and dimensions using volumetric distension of a catheter-mounted balloon and as an adjunctive test for the evaluation of symptoms suggestive of motor dysfunction. Targeted utilization of esophageal physiologic tests can lead to definitive diagnosis of esophageal disorders but can also help rule out organic disorders while making a diagnosis of functional esophageal disorders. Esophageal physiologic tests can evaluate obstructive symptoms (dysphagia and regurgitation), typical and atypical GERD symptoms, and behavioral symptoms (belching and rumination). Certain parameters from esophageal physiologic tests can help guide the management of GERD and predict outcomes. In this ACG clinical guideline, we used the Grading of Recommendations Assessment, Development and Evaluation process to describe performance characteristics and clinical value of esophageal physiologic tests and provide recommendations for their utilization in routine clinical practice.

Am J Gastroenterol 2020;00:1–17. https://doi.org/10.14309/ajg.0000000000000734; published online August 5, 2020

INTRODUCTION

Esophageal symptoms can arise from various esophageal disorders, from gastroesophageal reflux disease (GERD) to esophageal dysmotility to functional disorders. Esophageal physiologic tests are used not only to diagnose esophageal disorders but also to exclude obstructive motor disorders while diagnosing GERD or functional esophageal disorders and to identify dysmotility or behavioral disorders mimicking GERD. Symptoms of dysphagia and regurgitation can imply obstructive physiology in the esophagus, whereas heartburn, regurgitation, and chest pain may indicate the presence of GERD. Extraesophageal symptoms of cough, hoarseness, and globus are sometimes evaluated in the context of reflux disease. Atypical symptoms such as supragastric belching and rumination can mimic reflux. Consequently, esophageal physiologic testing has an important role in the clinical evaluation and management of esophageal disorders. However, it is important to emphasize that no test should be performed without a proper clinical history and without understanding of what the test will provide toward the patient’s diagnosis and/or management. To this end, the practitioner needs to understand the performance characteristics and clinical value of commonly used esophageal physiologic tests. This guideline summarizes the evidence underlying the use of each physiologic test and provides key concepts and recommendations for appropriate use of these tests.

METHODS

This guideline is structured in the format of summaries of evidence, recommendations, and key concepts pertaining to esophageal physiologic tests used for the evaluation of 3 symptom categories: obstructive symptoms, esophageal reflux symptoms, and extraesophageal or atypical symptoms. The authors of this manuscript developed specific patient population, intervention, comparator, and outcome (PICO) questions within the 3 symptom categories. A dedicated informationist (librarian) at the University of Michigan then performed literature searches to extract relevant manuscripts within the context of each PICO question. Two authors assigned to each PICO question each reviewed the literature searches and then concurred on supportive evidence specific to each question.

\textsuperscript{1}Division of Gastroenterology, Washington University School of Medicine, St. Louis, Missouri, USA; \textsuperscript{2}Division of Gastroenterology and Hepatology, Northwestern University Feinberg School of Medicine, Chicago, Illinois, USA; \textsuperscript{3}Department of Gastroenterology and Hepatology, University of Michigan, Ann Arbor, Michigan, USA; \textsuperscript{4}Division of Gastroenterology, Duke University & Durham VA Medical Center, Durham, North Carolina, USA; \textsuperscript{5}Department of Internal Medicine, Alameda Health System—Highland Hospital, Oakland, California, USA; \textsuperscript{6}Division of Gastroenterology and Hepatology, University of California in San Diego, La Jolla, California, USA. Correspondence: C. Prakash Gyawali, MD. E-mail: cprakash@dom.wustl.edu.

Received October 5, 2019; accepted May 4, 2020
quality evidence indicates very little confidence in the estimate. “Moderate” quality evidence is associated with moderate confidence in the effect estimate, although further research could affect the confidence of the estimate. “Low” quality evidence indicates that further study is likely to have an important impact on the confidence in effect estimate and would likely change the estimate. “Very low” quality evidence indicates very little confidence in effect estimate; the true effect is likely to be substantially different than the estimate of effect. A “strong” recommendation is made when the benefits clearly outweigh the negatives and the result of no action. “Conditional” is used when some uncertainty remains about the balance of benefits/potential harm. Key concepts are statements that are not amenable to the GRADE process, either because of the structure of the statement or because of the available evidence. In some instances, key concepts are based on extrapolation of evidence and/or expert opinion. Tables 2–5 summarize the GRADE recommendations and key concept statements in this guideline. Each recommendation statement has an associated assessment of the quality of evidence and strength of recommendation based on the GRADE process. Strengths of recommendations are not always contingent on the GRADE quality of evidence, particularly when the population health benefits are obvious and/or there is a suspected large magnitude of effect. Finally, the evidence summary for each section provides important definitions and data supporting the recommendations.

The Grading of Recommendations Assessment, Development and Evaluation (GRADE) system was used to evaluate the quality of supporting evidence (Table 1), with the GRADE process of evaluating quality of supporting evidence conducted by 2 formally trained GRADE methodologists (R.J.W. and R.Y.). The quality of the evidence is graded from high to very low. “High” quality evidence indicates that further research is unlikely to change confidence in the estimate of effect and that the true effect lies close to this estimate. “Moderate” quality evidence is associated with moderate confidence in the effect estimate, although further research could affect the confidence of the estimate. “Low” quality evidence indicates that further study is likely to have an important impact on the confidence in effect estimate and would likely change the estimate. “Very low” quality evidence indicates very little confidence in effect estimate; the true effect is likely to be substantially different than the estimate of effect. A “strong” recommendation is made when the benefits clearly outweigh the negatives and the result of no action. “Conditional” is used when some uncertainty remains about the balance of benefits/potential harm. Key concepts are statements that are not amenable to the GRADE process, either because of the structure of the statement or because of the available evidence. In some instances, key concepts are based on extrapolation of evidence and/or expert opinion. Tables 2–5 summarize the GRADE recommendations and key concept statements in this guideline. Each recommendation statement has an associated assessment of the quality of evidence and strength of recommendation based on the GRADE process. Strengths of recommendations are not always contingent on the GRADE quality of evidence, particularly when the population health benefits are obvious and/or there is a suspected large magnitude of effect. Finally, the evidence summary for each section provides important definitions and data supporting the recommendations.

**Table 1. Grading of Recommendations Assessment, Development and Evaluation quality criteria**

<table>
<thead>
<tr>
<th>Quality of evidence</th>
<th>Study design</th>
<th>Factors lowering the quality of evidence</th>
<th>Factors increasing the quality of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>Randomized trial</td>
<td>−1: serious limitations</td>
<td>Strong association:</td>
</tr>
<tr>
<td>Moderate</td>
<td>Observational study</td>
<td>−1: important inconsistency</td>
<td>+1: strong, no plausible confounders</td>
</tr>
<tr>
<td>Low</td>
<td>Any other evidence</td>
<td>Directness:</td>
<td>+1: very strong, no major threats to validity</td>
</tr>
<tr>
<td>Very low</td>
<td></td>
<td>−1: some uncertainty</td>
<td>+1: evidence of a dose response gradient</td>
</tr>
<tr>
<td></td>
<td></td>
<td>−2: major uncertainty</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>−1: sparse or imprecise data</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>−1: high probability of reporting bias</td>
<td></td>
</tr>
</tbody>
</table>

### Questionnaires

Multiple patient-report tools are available for standardized dysphagia evaluation, including the Eckardt symptom score, the Mayo Dysphagia Questionnaire, Brief Esophageal Dysphagia Questionnaire (BEDQ), and Eosinophilic Esophagitis Activity Index (1–4). The Mayo Dysphagia Questionnaire and BEDQ modestly differ between patients with and without major esophageal motor disorders (5,6). However, the BEDQ was only 70% sensitive and 65% specific in identifying major motor disorders (6) and neither differentiate between specific esophageal motor disorders (5,6). In eosinophilic esophagitis (EoE), a low symptom score (eosinophilic esophagitis activity index) was only able to detect histologic or endoscopic remission with approximately 60% accuracy (7). Although assessing dysphagia using standardized and validated patient-reported tools is an important practice to aid patient evaluation and to track outcomes, the inconsistent association with objective esophageal findings severely limits their utility to independently diagnose specific esophageal disorders (3,5,7).

### Key concepts

1. Patient-reported symptom questionnaires may aid the evaluation of patients with obstructive esophageal symptoms, but symptom questionnaires alone should not be used to diagnose specific esophageal conditions.

### Esophageal manometry

Esophageal manometry is generally considered the gold standard for the diagnosis of motility disorders. An alternate test is barium esophagography, which is available at most institutions, and has the capability to suggest the presence of motor disorders, including achalasia, and demonstrate anatomic relationships at the esophagogastric junction (EGJ). A retrospective study evaluating 281 patients at a single center with esophagram completed within 90 days of high-resolution manometry (HRM) demonstrated...

## Obstructive Symptoms

Several diagnostic procedures can be helpful in the evaluation of esophageal obstructive symptoms of dysphagia and regurgitation of esophageal contents. Evaluation starts with a careful history complemented with patient report tools. Core esophageal physiologic tests include esophagoscopy, manometry, and barium esophagography (Figure 1). Esophagoscopy has a very low yield as a test for esophageal physiology and motor pathophysiology but performs an essential role in excluding structural or mechanical obstructive lesions in the esophagus. Therefore, esophagoscopy during upper endoscopy, with biopsies, is an important first step in evaluating obstructive symptoms and should be performed before ordering esophageal physiologic studies.
significant disagreement between the 2 studies ($P = 0.04$) (8). The sensitivity and specificity of esophagram for detecting esophageal dysmotility were 0.69 and 0.50, respectively, with suboptimal positive and negative predictive values (0.61 and 0.58, respectively) (8). Esophagram therefore is a suboptimal screening examination for the detection of esophageal dysmotility in patients with esophageal dysphagia.

**Recommendations**

1. We recommend that patients with obstructive esophageal symptoms without a mechanical cause should undergo HRM for evaluation of esophageal motility disorders (conditional recommendation, very low quality of evidence).

Conventional manometry (CM) consists of stacked line tracings from pressure recordings extracted from widely spaced sensors mounted on a catheter. HRM represents an enhancement over CM in that pressure data are acquired from closely spaced circumferential sensors, which is then assimilated and displayed as 3 dimensional pressure topographs using dedicated software. Software tools interrogate pressure data to improve diagnostic accuracy. A multicenter randomized trial of 247 patients demonstrated an improved diagnostic yield for achalasia with HRM compared with CM (9). Diagnoses tended to be more frequently confirmed in patients who underwent HRM compared with CM (9). Furthermore, HRM provided superior inter-rater agreement and diagnostic accuracy for esophageal motility disorders compared with CM (10), and HRM-based subtyping of achalasia for which a CM-based paradigm does not exist, predicted the treatment outcome (11,12). Novice and intermediate learners demonstrated higher accuracy and reported greater ease at identification of obstructive motor disorders with HRM compared with CM (13).

The standard protocol for esophageal manometry involves 10 supine test swallows. Provocative maneuvers during HRM can augment information obtained from the 10 supine test swallow protocol and improve the diagnosis of motor disorders. Multiple rapid swallows consist of 5 repetitive 2 mL water swallows less than 3 seconds apart (14,15). During repetitive swallowing, esophageal contraction ceases and the lower esophageal sphincter (LES) relaxes. After the final swallow, robust esophageal body contraction occurs, termed “contraction reserve,” when contraction vigor exceeds that seen with standard single swallows. Absence of contraction reserve may be associated with a higher likelihood of postfundoplication dysphagia after antireflux surgery (ARS) (15,16) and development or worsening of ineffective esophageal motility over time (17). Rapid drink challenge consists of free drinking of 100–200 mL of water through a straw as fast as possible in the upright position and can aid identification of EGJ outflow obstruction via elevated LES postswallow residual pressures or panesophageal pressurization (18). A standardized test meal (typically consisting of cooked rice and gravy or a cheese and onion pasty) during HRM increases the diagnostic yield for obstructive motility disorders (EGJ outflow obstruction and spasm) and benefits interpretation by reproducing esophageal symptoms (19). In studies evaluating provocative maneuvers, there is uniform demonstration of added clinically useful information from provocative maneuvers compared with the interpretation of the 10-swallow protocol alone (15,18–22).

Table 2. Strength of the GRADE recommendations for esophageal physiologic testing for obstructive symptoms

<table>
<thead>
<tr>
<th>Statement</th>
<th>GRADE quality</th>
<th>Strength of recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>We suggest that patients with obstructive esophageal symptoms without a mechanical cause should undergo HRM for evaluation of esophageal motility disorders</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We recommend HRM over conventional line tracing manometry for the diagnosis of esophageal motility disorders in patients with obstructive esophageal symptoms</td>
<td>Moderate</td>
<td>Strong</td>
</tr>
<tr>
<td>We suggest utilization of supplementary/provocative maneuvers with the HRM protocol to improve the diagnostic yield of esophageal motility disorders in patients with obstructive esophageal symptoms</td>
<td>Low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the inclusion of a barium tablet with a barium esophagram during the evaluation of obstructive esophageal symptoms</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of FLIP to complement HRM for the diagnosis of esophageal motility disorders in patients with obstructive esophageal symptoms and borderline HRM findings.</td>
<td>Low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest that the EGJ and gastric cardia anatomy should be inspected endoscopically and/or radiographically to assess for mechanical abnormalities in patients with esophageal symptoms after antireflux surgery</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
</tbody>
</table>

Please refer to the text for detailed discussion and evidence regarding each of the recommendations.

EGJ, esophagogastric junction; FLIP, functional lumen imaging probe; GRADE, Grading of Recommendations Assessment, Development and Evaluation; HRM, high-resolution manometry.
Table 3. Strength of the GRADE recommendations for esophageal physiologic testing for typical reflux symptoms

<table>
<thead>
<tr>
<th>Statement</th>
<th>GRADE quality</th>
<th>Strength of recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>We suggest the use of ambulatory reflux monitoring over patient-reported symptoms on GERD questionnaires for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of ambulatory reflux monitoring over the assessment of response to PPI therapy for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of ambulatory reflux monitoring over upper endoscopy alone (if endoscopy is not definitive) for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms not responding to PPI</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of ambulatory reflux monitoring performed off PPI therapy over ambulatory reflux monitoring on PPI therapy for a conclusive diagnosis of GERD in patients with typical reflux symptoms and unproven GERD</td>
<td>Low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of prolonged wireless pH monitoring over 24-h catheter-based monitoring for the diagnosis of GERD in adults with infrequent or day-to-day variation in esophageal symptoms (conditional recommendation, very low quality of evidence).</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest the use of ambulatory pH impedance monitoring on PPI therapy over endoscopic evaluation or pH monitoring alone to diagnose persisting GERD in adults with typical esophageal reflux symptoms and previous confirmatory evidence of GERD (proven GERD)</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest that for patients with esophageal symptoms being evaluated for antireflux surgery, abnormal acid exposure time be considered a predictor of treatment outcome; reflux symptom association and mean nocturnal baseline impedance provide adjunctive value</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
</tbody>
</table>

Please refer to the text for detailed discussion and evidence regarding each of the recommendations.

GERD, gastroesophageal reflux disease; GRADE, Grading of Recommendations Assessment, Development and Evaluation; PPI, proton pump inhibitor.

Table 4. Strength of the GRADE recommendations for esophageal physiologic testing for extraesophageal reflux symptoms and atypical symptoms

<table>
<thead>
<tr>
<th>Statement</th>
<th>GRADE quality</th>
<th>Strength of recommendation</th>
</tr>
</thead>
<tbody>
<tr>
<td>We recommend ambulatory reflux monitoring, specifically pH impedance monitoring performed off acid suppression, over laryngoscopy for a diagnosis of extraesophageal reflux</td>
<td>Low</td>
<td>Strong</td>
</tr>
<tr>
<td>We suggest up-front ambulatory reflux monitoring off acid suppression over an empiric trial of PPI therapy for extraesophageal reflux symptoms without concurrent typical reflux symptoms</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest HRIM with postprandial monitoring be used to confirm the diagnosis of rumination if clinically necessary in patients with esophageal symptoms suspicious for rumination syndrome</td>
<td>Low</td>
<td>Conditional</td>
</tr>
<tr>
<td>We suggest that for patients with excessive belching, pH impedance monitoring can be used to confirm the diagnosis of supragastric belching</td>
<td>Very low</td>
<td>Conditional</td>
</tr>
</tbody>
</table>

Please refer to the text for detailed discussion and evidence regarding each of the recommendations.

GRADE, Grading of Recommendations Assessment, Development and Evaluation; HRIM, high-resolution impedance manometry; PPI, proton pump inhibitor.

Recommendations

3. We recommend the utilization of supplementary/provocative maneuvers with the manometry protocol to improve the diagnostic yield of esophageal motility disorders in patients with obstructive esophageal symptoms (conditional recommendation, low quality of evidence).

Barium esophagram

Barium esophagram can assess esophageal bolus transit. Although multiple studies report the utility of a timed upright barium esophagram in evaluating achalasia, particularly achalasia outcomes (23–26), there are no studies that directly compare timed barium esophagram with nontimed barium esophagram. Normative values are also not available, and reported protocols are somewhat variable (use of 100–250 mL thin barium). Nevertheless, a timed upright esophagram performed using 8 oz or 236 mL barium, evaluating barium height at 1 minute (abnormal when >5 cm) and 5 minutes (abnormal >2 cm) provides evidence for abnormal esophageal emptying, not just in achalasia but also in other esophageal outflow obstruction syndromes (27).

Key concepts

2. When performing an esophagram for the evaluation of patients with obstructive esophageal symptoms, a standardized, upright, timed barium esophagram protocol should be used.
In a study evaluating combined liquid barium and a 13-mm barium tablet to liquid barium alone, both liquid barium and barium tablet transit were abnormal more often (74.8%) in 107 patients with untreated achalasia; barium tablet transit was abnormal with normal liquid barium transit in 48.9% of 45 patients with EGJ outflow obstruction, and both were normal in 60.6% of 132 patients without achalasia, with statistically significant differences between the groups (27). Abnormal passage or retention of a 13-mm barium tablet can thus be indicative of an obstructive process at the EGJ (27,28).

(Continued)

Table 5. Key concepts

<table>
<thead>
<tr>
<th>Key concepts</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient-reported symptom questionnaires may aid evaluation of patients with obstructive esophageal symptoms, but symptom questionnaires alone should not be used to diagnose specific esophageal conditions.</td>
</tr>
<tr>
<td>When performing an esophagram for evaluation of patients with obstructive esophageal symptoms, a standardized, upright, timed barium esophagram protocol should be used.</td>
</tr>
<tr>
<td>Barium esophagography provides information about bolus clearance in patients with dysphagia; HRIM provides adjunctive information about bolus clearance.</td>
</tr>
<tr>
<td>In patients whom a manometry study cannot be completed, such as catheter placement failure despite attempts at endoscopic placement, FLIP topography may be used for the diagnosis of esophageal motility disorders.</td>
</tr>
<tr>
<td>When available, FLIP can be used to measure EGJ distensibility or minimal EGJ cross-sectional area intraprocedurally during invasive treatment for achalasia.</td>
</tr>
<tr>
<td>When available, FLIP may be considered for the measurement of distensibility to assess the fibrostenotic remodeling of the esophagus and stratify risk of food impaction in patients with eosinophilic esophagitis.</td>
</tr>
<tr>
<td>Patient-reported outcome measurement during follow-up after treatment in achalasia, accompanied by an objective measure of esophageal function (e.g., timed barium esophagram) may be used to assess the treatment outcome.</td>
</tr>
<tr>
<td>Esophageal HRM complements the diagnostic evaluation of chest pain not responsive to PPI therapy.</td>
</tr>
<tr>
<td>HRM is important for ruling out motor disorders and for assessing esophageal peristaltic performance in patients with GERD symptoms; HRM should be performed before antireflux surgery or invasive GERD management.</td>
</tr>
<tr>
<td>HRM complements endoscopy and barium esophagography in improving the diagnostic yield of identifying hiatal hernia.</td>
</tr>
<tr>
<td>Please refer to the text for detailed discussion and evidence regarding each of the key concepts.</td>
</tr>
</tbody>
</table>

© 2020 by The American College of Gastroenterology

The American Journal of GASTROENTEROLOGY

In a study evaluating combined liquid barium and a 13-mm barium tablet to liquid barium alone, both liquid barium and barium tablet transit were abnormal more often (74.8%) in 107 patients with untreated achalasia; barium tablet transit was abnormal with normal liquid barium transit in 48.9% of 45 patients with EGJ outflow obstruction, and both were normal in 60.6% of 132 patients without achalasia, with statistically significant differences between the groups (27). Abnormal passage or retention of a 13-mm barium tablet can thus be indicative of an obstructive process at the EGJ (27,28).

Recommendations

4. We recommend inclusion of a barium tablet with a barium esophagram during the evaluation of obstructive esophageal symptoms (conditional recommendation, very low quality of evidence).

With the availability of intraluminal impedance measurements during HRM, liquid bolus clearance can be assessed using high-resolution impedance manometry (HRIM). In a study comparing bolus transit between HRIM and barium esophagram in 20 patients with achalasia, impedance-barium esophagram concordance was found to be high for swallows with normal esophageal emptying and for severe barium stasis (29).

Key concepts

3. Barium esophagography provides information about bolus clearance in patients with dysphagia; HRIM provides adjunctive information about bolus clearance.

Functional lumen imaging probe

The functional lumen imaging probe (FLIP) is a Food and Drug Administration-approved measurement tool used to measure simultaneous pressure, cross-sectional area (CSA), and distensibility in the esophagus. Although commercially available since 2009, FLIP has limited penetration into clinical settings outside of specialized centers because of a lack of standardized protocols, lack of data analysis methodology, and paucity of data supporting utility in general practice. In a study comparing esophageal motility assessed using FLIP topography to HRM in patients with dysphagia, FLIP was well-tolerated and accurately detected major motility disorders including achalasia (30). FLIP topography enhanced the evaluation of esophageal function in nonobstructive dysphagia by detecting an abnormal response to esophageal distension in 50% of patients diagnosed with ineffective esophageal motility or a normal HRM study (31). Furthermore, FLIP can characterize achalasia subtypes by detecting nonocclusive esophageal contractions not observed with HRM. Such contractility was detected to varying degrees in each of the achalasia subtypes, potentially allowing additional subclassification of patients with achalasia (32). EGJ distensibility measured using FLIP can diagnose achalasia in patients with clinically suspected achalasia but manometrically normal EGJ relaxation (33), a known entity that represents a caveat for the use of integrated relaxation pressure alone in excluding achalasia (34,35). Thus, FLIP can identify an obstructive element in major motor disorders presenting with dysphagia despite a normal integrated relaxation pressure but is not intended to replace HRM in the characterization of motor disorders (30,33). The value of FLIP lies in the identification of achalasia or esophageal outflow obstruction in patients with borderline manometric findings or in patients with obstructive symptoms despite the management of esophageal outflow obstruction (Figure 1). However, more research is needed before this technology can replace conventional means of esophageal testing, and studies regarding use of FLIP as an adjunct to existing esophageal tests needs validation from independent researchers with no real or perceived bias.
We recommend the use of FLIP to complement HRM for the diagnosis of esophageal motility disorders in patients with obstructive esophageal symptoms and borderline HRM findings (conditional recommendation, low quality of evidence). In patients whom a manometry study cannot be completed, such as catheter placement failure despite attempts at endoscopic placement, FLIP topography may be used for the diagnosis of esophageal motility disorders.

FLIP can direct invasive achalasia treatments and can predict clinical outcomes (36,37). Studies evaluating intraoperative CSA measurements demonstrated correlation of the final EGJ CSA during per oral endoscopic myotomy (38) and surgical myotomy (39,40) for achalasia with clinical response. Other investigators similarly report an increase in the EGJ diameter and distensibility index after per oral endoscopic myotomy (41). Intraprocedural EGJ distensibility measurements correlate with immediate symptom outcome after pneumatic dilation (42). However, the exact FLIP protocol to be used and target values for post-treatment distensibility and CSA have not been defined and further research is needed.

**Key concepts**

4. We recommend the use of FLIP to complement HRM for the diagnosis of esophageal motility disorders in patients with obstructive esophageal symptoms and borderline HRM findings (conditional recommendation, low quality of evidence). In patients whom a manometry study cannot be completed, such as catheter placement failure despite attempts at endoscopic placement, FLIP topography may be used for the diagnosis of esophageal motility disorders.

FLIP has also been studied in EoE, where patients with previous food impaction had significantly lower distensibility plateau values than those with solid food dysphagia alone (43). Reduced esophageal distensibility in EoE may predict the risk for food impaction and indicate the requirement for esophageal endoscopic myotomy.
dilation in EoE. FLIP has been demonstrated to be feasible and useful as a marker for esophageal remodeling in both pediatric and adult EoE populations (44,45). However, further research is needed before this indication for FLIP can become a part of routine clinical care in EoE.

**Key concepts**

6. When available, FLIP may be considered for measurement of distensibility to assess fibrostenotic remodeling of the esophagus and stratify risk of food impaction in patients with EoE.

**Testing after achalasia management**

Improvement and possible resolution of patient symptoms is an important goal of treatment, and thus, measurement of patient-reported outcomes (PROs) is useful during the follow-up after achalasia treatment. The Eckardt score is a simple and commonly used questionnaire that semiquantitates severity of 4 items: dysphagia, regurgitation, chest pain, and weight loss (1). However, it was developed without rigorous evaluation for validity and reliability and subsequent analysis suggests only fair psychometric performance, with particular weakness related to the chest pain and weight loss items (46). In addition, discordance is sometimes observed between symptom severity and objective esophageal function (such as esophageal retention quantified with timed barium esophagram) after the treatment of achalasia (23–26). Furthermore, objective esophageal retention on a timed upright barium esophagram may be a better predictor for future treatment failure and need for retreatment in achalasia (23,25). However, available data do not provide direction on the use of PROs vs objective testing after achalasia therapy and whether either mode of post-treatment evaluation can be used in lieu of the other.

**Key concepts**

7. PRO measurement during the follow-up after treatment in achalasia, accompanied by an objective measure of esophageal function (e.g., timed barium esophagram) may be used to assess the treatment outcome.

**TYPICAL REFLUX SYMPTOMS**

Evaluation starts with a careful history. When patients present with symptoms of heartburn and acid regurgitation, an empiric trial of acid suppressive therapy is typically used; this approach may also be used for chest pain presentations where a cardiac source has been ruled out. Although this is adequate for initial management, neither symptom assessment (GERD questionnaires) nor response to proton pump inhibitor (PPI) trials are adequate for conclusive diagnosis of GERD, which is necessary before invasive management of GERD. The standard for assessment of abnormal esophageal acid exposure is ambulatory reflux monitoring, either pH monitoring or pH impedance monitoring (Figure 1). This may not be necessary if endoscopy demonstrates high-grade erosive esophagitis or evidence of GERD-related esophageal complications (Barrett’s esophagus, peptic stricture). Esophageal HRM may demonstrate pathophysiologic mechanisms underlying GERD and is emerging as an adjunctive method of value when evidence for GERD is otherwise inconclusive.

**GERD questionnaires**

GERD questionnaires can standardize reporting of reflux symptoms, but these do not necessarily correspond to pathologic GERD on objective testing. Among 85 patients undergoing 24-hour pH impedance monitoring, the six-item GERDQ score ≥8 had 100% sensitivity but 37% specificity for acid exposure time (AET) > 6.3% off PPI and 75% sensitivity but 26% specificity on PPI (47). The Diamond study with a broader definition of GERD (including abnormal AET and positive symptom response to PPI) found that the 12-item reflux disease questionnaire (a precursor of the GERDQ) had 62% sensitivity and 67% specificity for GERD (48). A multicenter study of 169 Norwegian patients found that GERDQ scores ≥9 had 66% sensitivity and 64% specificity for GERD defined as any of reflux esophagitis on esophagogastroduodenoscopy (EGD), total AET ≥5.5%, supine AET ≥6.9%, upright AET ≥6.7%, or positive symptom association probability (SAP) (49). The Mayo-GERD questionnaire when compared with distal esophageal AET >4% on ambulatory reflux monitoring in a cohort of over 300 patients had a sensitivity of 68% and specificity of 72% at the optimal threshold (50). These findings suggest that GERD questionnaires have modest performance characteristics for a conclusive GERD diagnosis.

**Recommendations**

5. We recommend the use of ambulatory reflux monitoring over patient-reported symptoms on GERD questionnaires for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms (conditional recommendation, very low quality of evidence).

**Empiric PPI trial**

An empiric trial of PPI (the “PPI test”) is a pragmatic approach to typical reflux symptoms in clinical practice, given limited invasiveness, lower cost, and symptomatic response corroborating clinical suspicion for GERD. The original PPI test consisted of 40–80 mg a day of omeprazole or equivalent, typically in divided doses for 7–28 days (51,52), but various modifications have been used in clinical studies. A meta-analysis of 15 studies evaluating empiric PPI trials (of 1–4 weeks in duration) against ambulatory pH testing as the reference standard demonstrated a sensitivity of 78% but specificity of only 54% for a diagnosis of GERD (53). Meta-analyses of studies evaluating empiric PPI therapy for noncardiac chest pain (NCCP) (with erosive esophagitis and/or 24-hour pH monitoring as reference standards) also found 80% sensitivity for this approach (54,55).

In an analysis of data from the Diamond study, when a 2-week PPI test was compared against the presence of any of reflux esophagitis at EGD, abnormal AET >5.5%, or positive SAP >95%, 69% of patients with GERD had a positive response compared with 51% without GERD, indicating a positive likelihood ratio of 1.52 and a negative likelihood ratio of 0.71 (48,56). These data reinforce the limited diagnostic utility of the PPI test to conclusively identify patients with GERD. The clearest need for objective reflux monitoring for a conclusive GERD diagnosis is in symptomatic patients who do not respond to acid suppressive therapy, patients on whom invasive reflux management is planned, and patients concerned about long-term PPI therapy.
Recommendations

6. We recommend the use of ambulatory reflux monitoring over the assessment of response to PPI therapy for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms (conditional recommendation, very low quality of evidence).

Endoscopy

Endoscopy provides an important role for the evaluation of reflux symptoms to objectively diagnose reflux in the presence of high-grade erosive esophagitis, Barrett’s esophagus, or peptic stricture. However, the increasing popularity of PPI trials for suspected GERD symptoms has decreased the likelihood of finding reflux esophagitis on EGD. Among 696 patients undergoing EGD for suspected GERD symptoms, those without reflux esophagitis were more likely to be on PPI therapy compared with those with erosive esophagitis (53% vs 29%, univariate odds ratio 2.75, P < 0.001; multivariate odds ratio 3.19, P < 0.001) (57). Among over 700 patients with a partial response to PPI therapy, only 20–30% had esophageal mucosal breaks on EGD (58). Consequently, despite high specificity, EGD has low sensitivity for a diagnosis of GERD, but a good quality EGD is essential before embarking on further evaluation of esophageal symptoms.

Recommendations

7. We recommend the use of ambulatory reflux monitoring over upper endoscopy alone (if endoscopy is not definitive) for a conclusive diagnosis of GERD in patients with esophageal reflux symptoms not responding to PPI (conditional recommendation, very low quality of evidence).

Ambulatory reflux monitoring

Symptoms, PPI response, and low-grade erosive esophagitis (Los Angeles, LA grades A and B) on endoscopy are insufficient conclusive evidence for GERD and do not always correlate with abnormal reflux burden on ambulatory pH monitoring performed off PPI therapy (59). Hence, these constitute unproven GERD and require conclusive evidence of GERD before escalation of management, with reflux testing performed off antisecretory therapy. Most patients not responding to PPI therapy (60) and as many as half of patients referred for invasive GERD management will not have pathologic reflux evidence on ambulatory reflux monitoring (61), which is significant because pathologic reflux burden, particularly abnormal AET, predicts symptom improvement from antireflux therapy, including surgery (60,62). In a retrospective analysis of prospectively collected data in 188 patients studied on and off PPI therapy before definitive antireflux management predictors of symptom improvement on a multivariate analysis included AET, reflux symptom association (RSA, includes symptom index and SAP), and testing off PPI therapy for 7 days (62). By contrast, in a retrospective analysis of 33 patients undergoing pH impedance monitoring on PPI therapy before fundoplication, the only predictor of symptom improvement was RSA (63). AET and RSA from ambulatory reflux monitoring can phenotype GERD presentations into strong evidence for GERD (abnormal AET, positive RSA), good evidence (abnormal AET, negative RSA), reflux hypersensitivity (normal AET, positive RSA), or no evidence (normal AET, negative RSA) (64,65); these phenotypes can stratify the symptomatic outcome from medical or surgical antireflux therapy (64). However, RSA alone has modest performance characteristics in predicting reflux management outcome (66) and is subject to accuracy of symptom reporting (67); therefore, RSA is best used as an adjunctive measure to complement AET (65).

Recommendations

8. We recommend the use of ambulatory reflux monitoring performed off antisecretory therapy over ambulatory reflux monitoring on antisecretory therapy for a conclusive diagnosis of GERD in patients with typical reflux symptoms and unproven GERD (conditional recommendation, low quality of evidence).

Prolonged wireless pH monitoring can highlight day-to-day variation in esophageal acid exposure and augment the diagnosis of pathologic GERD even when the first 24 hours of a multiday study is negative for GERD. In a cohort of 471 patients evaluated with prolonged wireless pH monitoring off PPI, 56% of patients with heartburn had abnormal AET >5.5%; the presence of a hiatus hernia and body mass index >25 were predictors of abnormal AET (68). Using wireless pH monitoring, extended recording time of 48–96 hours increases the diagnostic yield, both for increased identification of abnormal reflux burden and for RSA (69–71). Diagnosis may shift to NERD from functional heartburn if additional days’ data are taken into consideration, in comparison to day 1 data (72). Wireless pH monitoring is particularly useful when the transnasal catheter is not tolerated or yields a negative result despite high suspicion of GERD (72). However, wireless pH monitoring is expensive, limiting its availability in many countries.

Recommendations

9. We recommend the use of prolonged wireless pH monitoring over 24-hour catheter-based monitoring for the diagnosis of GERD in adults with infrequent symptoms or day-to-day variation in esophageal symptoms (conditional recommendation, very low quality of evidence).

Advanced grade erosive esophagitis and confirmed Barrett’s esophagus constitute conclusive evidence for GERD; consequently, pH impedance metrics in these settings are consistently abnormal off PPI therapy (73,74). Hence, advanced grade erosive esophagitis and confirmed Barrett’s esophagus are considered proven GERD, where reflux testing can be performed on therapy. For treatment of persisting symptoms in patients with proven GERD on maximal (twice a day) PPI therapy, expert esophagologists recommend invasive therapy only in the presence of conclusive evidence of GERD, including abnormal reflux burden with or without hiatal hernia or regurgitation with positive symptom-reflux association and a large hiatus hernia (75). GERD can be established in this setting by performing pH impedance monitoring on submaximal or maximal PPI therapy. When 39 patients with refractory reflux symptoms were tested both on therapy (pH impedance monitoring) and off therapy (wireless pH monitoring), weakly acid reflux episodes were more frequently encountered on therapy in patients with abnormal AET off therapy, reinforcing the value of pH impedance monitoring on therapy in proven GERD (76). Abnormal pH impedance on once a day PPI therapy normalized with maximal PPI therapy in 71.1% of 45 patients (77). Furthermore, 89% of 38 patients with refractory symptoms and abnormal pH impedance metrics on maximal PPI therapy improved with laparoscopic fundoplication. These data suggest that escalation of reflux management can benefit patients with proven GERD who continue to have abnormal pH.
impedance metrics on submaximal or maximal medical antireflux therapy. In addition, the use of pH impedance monitoring over pH monitoring alone shifts diagnoses from functional heartburn to reflux hypersensitivity because more reflux episodes are identified using pH impedance over pH monitoring alone, and the symptoms may associate with reflux episodes detected by pH impedance (78). Patients with proven GERD and persisting symptoms despite antireflux therapy are anticipated to form a very small fraction of overall patients with GERD; some may require ambulatory reflux monitoring on PPI (perhaps in addition to initial testing off PPI) to decide if invasive GERD management is indicated for persisting symptoms.

Recommendations

10. We recommend the use of ambulatory pH impedance monitoring on PPI therapy over endoscopic evaluation or pH monitoring alone to diagnose persisting GERD in adults with typical esophageal reflux symptoms and previous confirmatory evidence of GERD (proven GERD) (conditional recommendation, very low quality of evidence).

Esophageal manometry

Once cardiac etiologies have been appropriately excluded, GERD is the most common mechanism for NCCP, retrosternal angina like chest pain without a cardiac cause (79). In support of this, meta-analyses suggest that a trial of empiric PPI therapy has 80% sensitivity for a GERD diagnosis (54,55). However, if chest pain does not respond to PPI therapy, esophageal HRM is an important diagnostic modality because esophageal dysmotility, specifically achalasia, spasm, or hypercontractility can be an explanation for chest pain, albeit an epiphenomenon rather than a cause (80,81). For instance, among 177 patients with NCCP who underwent manometry and pH testing, 35% were diagnosed with GERD, whereas 7% had jackhammer esophagus, 5% had distal esophageal spasm, and 2% had achalasia (82). Furthermore, esophageal motor assessment has important implications for the management of GERD (83,84) and is required to make a diagnosis of functional chest pain (79).

Key concepts

8. Esophageal HRM complements the diagnostic evaluation of chest pain not responsive to PPI therapy.

Esophageal manometry is often performed as part of esophageal function testing to rule out esophageal motor disorders, localize the LES for ambulatory reflux catheter placement, and assist in preoperative planning for GERD (83,85,86). In one study, 3% of patients undergoing HRM before planned fundoplication had achalasia spectrum disorders; proceeding with standard fundoplication in these patients would have worsened the obstruction (87). A study of 524 patients with achalasia found that 29% had been treated unsuccessfully with antisecretory medications and referred for ARS (88); other studies similarly highlight the overlap between motor disorders and GERD symptoms (82,89,90). Finally, in carefully selected clinical scenarios, postprandial HRIM can identify ruminative syndrome, which has suboptimal outcomes with standard antireflux therapies (91). Therefore, HRM can diagnose motor disorders that can mimic GERD and can demonstrate the adequacy of esophageal peristalsis before invasive GERD therapy, although HRM by itself cannot diagnose GERD.

Key concepts

9. HRM is important for ruling out motor disorders and for assessing esophageal peristaltic performance in patients with GERD symptoms; HRM should be performed before ARS or invasive GERD management.

HRM provides morphological details of the EGJ. By comparing the relative locations of the intrinsic LES and the crural diaphragm, the presence of a hiatus hernia can be identified (83,92). In a study evaluating 215 patients with and without hiatus hernia, manometric hiatus hernia (using CM) was larger compared with endoscopy ($P < 0.001$). Against an endoscopic gold standard, CM had 28% sensitivity, 97% specificity, and 82% positive predictive value in identifying a sliding hiatal hernia (93). However, these data need to be interpreted with caution because CM was used, and endoscopic identification of a hiatus hernia has poor sensitivity despite high specificity when compared with barium esophagography. In a study of 92 obese subjects evaluated before bariatric surgery, the sensitivity of endoscopy for detection of sliding hiatus hernia was $\leq 40\%$ despite high specificity ($\geq 94\%$) compared with barium radiography (94). In addition, body position affects manometric detection of a hiatus hernia, with a higher detection rate in the upright or standing position compared with supine (95).

Comparisons between HRM and surgical identification of a hiatus hernia suggest higher sensitivity and accuracy with HRM. In a retrospective study of 83 consecutive patients undergoing laparoscopic fundoplication, 42 patients had a hiatus hernia $>2$ cm at surgery. False positive rates with preoperative HRM diagnosis of hiatus hernia (5%) were significantly lower compared with endoscopic diagnosis (32%, $P = 0.01$), whereas false negative rates were similar (48% vs 45%, $P = ns$) (96). In a prospective study of 34 obese patients undergoing bariatric surgery, HRM had better performance characteristics (sensitivity 88.9%, specificity 60.0%) compared with endoscopy or barium radiography (sensitivity 77.4, specificity 44.0%) using surgical identification of hiatus hernia as the gold standard (97). Another study comparing HRM, endoscopy, and barium radiography in 90 patients demonstrated a sensitivity of 92% and specificity of 95% with HRM compared with detection with the other 2 tests (98). Finally, in a prospective study of 100 consecutive patients, 53 of which had a hiatus hernia on surgical measurement during laparoscopy, preoperative HRM had a sensitivity of 94.3% and specificity of 91.5% compared with endoscopy (96.2%, 74.5%) or barium radiography (69.8%, 97.9%) (99).

Thus, available data suggest a higher sensitivity of HRM for hiatus hernia detection compared with either endoscopy or barium radiography. However, because of varying performance characteristics, the 3 studies are complementary.

Key concepts

10. HRM complements endoscopy and barium esophagography in improving the diagnostic yield of identifying hiatal hernia.

Management implications

Adequate preoperative evaluation and appropriate patient selection are critical to treatment success with invasive antireflux management, including ARS, and ambulatory reflux monitoring is important as part of this evaluation (100). In a retrospective
study of 62 PPI nonresponders with suspected nonerosive reflux disease, 66% had normal acid exposure on pH impedance testing off antisecretory therapy (101). In a prospective study of 366 patients with refractory heartburn who were enrolled in a Veterans Administration study, 99 (27%) had functional heartburn on the basis of negative esophageal testing including pH impedance monitoring on acid suppression, whereas 23 (6%) had non-GERD esophageal disorders and 7 (2%) had esophageal motility disorders (60). Similarly, in a retrospective study of 221 patients referred for ambulatory reflux monitoring, only 45% had confirmation of GERD (61). Thus, patients with typical GERD symptoms, normal EGD, and poor PPI response may have non-GERD etiologies for their symptoms and should not be referred for ARS (102).

Distal esophageal AET is a cardinal reflux metric that predicts GERD treatment outcome. In a retrospective study of 683 patients with suspected GERD, AET >4.0% and a positive symptom index both predicted PPI response (103). In a prospective study, 88% of patients with an objective diagnosis of GERD, with either erosive esophagitis on EGD or AET >4.2%, reported symptom relief with PPI therapy (104). In another study of 128 patients referred for pH impedance testing off PPI, AET >4.0% predicted symptom improvement with PPI therapy (105). Although many studies used AET thresholds of approximately 4.0% to designate GERD, the recent Lyon consensus proposes that AET >6% be considered pathologic and AET 4–6% be considered borderline with the need for additional GERD evidence to confirm pathologic GERD; these thresholds were based on existing normative data and expert opinion (65).

Similarly, pathologic AET and RSA also predict treatment success from ARS. In a study of 187 subjects referred for pH impedance monitoring before medical or surgical antireflux management, AET >4%, RSA with impedance-detected reflux events, and testing performed off PPI therapy predicted treatment success (62). In another study of 33 patients who underwent laparoscopic fundoplication, the only predictor of successful postoperative outcome was positive RSA on pH impedance performed on PPI therapy (63).

Baseline impedance (BI) has gained interest as a novel impedance metric with the ability to segregate GERD from non-GERD processes because BI correlates inversely with esophageal mucosal integrity and esophageal acid burden. Esophageal BI can be acquired from 24-hour pH impedance tracings as mean nocturnal baseline impedance (MNBI) (106). This is averaged from three 10-minute periods in the distal-most 2 impedance channels during the nocturnal sleep period, when there are limited artifacts, swallows, and reflux episodes (65,107–111); it is anticipated that automated analysis will be available in the future. BI can also be evaluated using balloon catheters as mucusal impedance (MI) (112) or from HRIM studies as BI-HRIM (113). Although previous work suggested that BI had a 78% sensitivity and 71% specificity for differentiating reflux disease from functional heartburn (114), in a larger cohort of PPI-responsive heartburn, an MNBI threshold of <2,292 Ω identified those with erosive reflux disease with 91% sensitivity and 86% specificity and those with pH-positive GERD with 86% sensitivity and 86% specificity (110). Furthermore, among a cohort of nearly 100 patients with a prospective follow-up of 3 years, univariate and multivariate analyses showed that distal MNBI was predictive of symptomatic improvement with medical or surgical antireflux therapy (115). In fact, a retrospective study of over 400 patients found that MNBI linked reflux with PPI responsiveness better than AET (107). Among patients for whom AET is inconclusive, abnormal MNBI values <2,292 Ω predict symptomatic response with medical or surgical antireflux therapy (116). However, further prospective studies and meta-analyses of existing studies are needed to better determine the precise role of BI including MNBI in clinical reflux evaluation and management.

After a reflux episode, a primary swallow, termed the postreflux swallow-induced peristaltic wave (PSPW), is often seen in healthy individuals, which serves to bring saliva for neutralization of esophageal mucosal acidification and is therefore a marker of esophageal chemical clearance (117). The proportion of reflux episodes with a PSPW within 30 seconds among the total number of reflux episodes on a pH impedance study constitutes the PSPW index. The PSPW index is lower in erosive or nonerosive GERD compared with functional heartburn or healthy controls (108,109). Data from a single research group suggest that PSPW measurements might outperform AET and MNBI in predicting the PPI responsiveness in endoscopy-negative heartburn (107,118). These data need to be replicated by other investigators, and further research is needed before widespread use of this novel impedance metric in clinical esophagology.

Although controversies remain regarding the role of pH impedance testing as opposed to pH testing alone and testing on or off PPI therapy in the preoperative evaluation for ARS, it is clear that some form of reflux documentation is essential before invasive GERD management. Among a cohort of nearly 100 patients who underwent pH impedance monitoring off PPI therapy, phenotyping patients based on (i) abnormal or physiologic AET and (ii) positive or negative RSA demonstrated that symptomatic outcomes with antireflux therapy were best with strong or good evidence for GERD on testing but poorest in the setting of physiologic AET and negative RSA (64). Cost modeling suggests that early referral for ambulatory reflux monitoring, as long as the sensitivity of pH monitoring remains above 35%, may be more cost-effective than the prolonged PPI trials often used in clinical practice (119) because early ambulatory testing may support averting PPI use in potentially half of tested patients (120). Conversely, negative ambulatory reflux monitoring studies (with physiologic distal esophageal AET and negative RSA) may suggest the presence of non-GERD processes (such as visceral hypersensitivity, esophageal motor disorders, behavioral disorders, and EoE) and predict poor symptomatic responses with antireflux treatments.

**Recommendations**

11. We recommend that for patients with esophageal symptoms being evaluated for antireflux surgery, abnormal AET be considered a predictor of treatment outcome; RSA and mean nocturnal BI provide adjunctive value (conditional recommendation, very low quality of evidence).

**Testing after ARS**

In patients who develop symptoms after ARS, disrupted integrity of the antireflux surgical site may have implications on management. Both endoscopy and esophagography can be used to assess integrity of the fundoplication and to identify slippage, displacement, and recurrence of a hiatus hernia (121,122). An intact fundoplication is associated with successful symptomatic outcome in 81.7% of patients (121). Normal radiologic and endoscopic evidence of an intact fundoplication correlates with normal manometry and
24-hour pH monitoring. However, there are little data comparing the diagnostic yield of esophagram vs endoscopy in detecting a defective fundoplication wrap (122).

**Recommendations**

12. We recommend that the EGJ and gastric cardia anatomy should be inspected endoscopically and/or radiographically to assess mechanical abnormalities in patients with esophageal symptoms after ARS (conditional recommendation, very low quality of evidence).

**EXTRAESOPHAGEAL AND ATYPICAL SYMPTOMS**

### Extraesophageal symptoms

Attributing extraesophageal symptoms to GERD has gained momentum since the 1990s and continues to increase (123). On average, patients with extraesophageal symptoms will visit 10 consultants and undergo 6 diagnostic procedures in the first year of evaluation, often without diagnostic clarity (124,125), contributing to more than $5,000 in annual health care costs per patient (125).

A diagnosis of laryngopharyngeal reflux (LPR) is often made after a laryngoscopy. However, laryngoscopic findings of erythema, edema, and/or postcricoid hyperplasia have low specificity for GERD and are common in healthy volunteers (126). In small prospective cohorts of patients with laryngesophageal symptoms and laryngoscopic signs of LPR, pH impedance testing off PPI therapy confirmed GERD in less than half of patients; prevalence of symptoms and laryngoscopic findings were similar regardless of positive or negative reflux monitoring (127–129). Performance characteristics of laryngoscopy for extraesophageal reflux compared with reflux monitoring consisted of 86% sensitivity, 9% specificity, and 44% diagnostic accuracy (128), and reflux finding scores from laryngoscopy did not correlate with pH impedance test findings (129). Similarly, in a prospective study of 33 patients with suspected LPR, laryngoscopic findings did not predict the response to PPI therapy (130). Furthermore, the inter-rater reliability and agreement between otolaryngologists for laryngoscopic findings suggestive of LPR is suboptimal (131).

Thus, although most data were derived from prospective cohort studies with small sample sizes, they all point to the lower specificity of laryngoscopy compared with ambulatory reflux monitoring for a diagnosis of extraesophageal reflux.

**Recommendations**

13. We recommend ambulatory reflux monitoring, specifically pH impedance monitoring performed off acid suppression, over laryngoscopy for a diagnosis of extraesophageal reflux (strong recommendation, low quality of evidence).

Although most patients with suspected extraesophageal reflux will undergo an empiric PPI trial, the results from meta-analyses examining this approach are mixed (132–134). One meta-analysis of 72 studies, including 10 randomized controlled trials, reported a relative risk of 1.31 in favor of PPIs for extraesophageal reflux, although the meta-analysis also highlighted the significant heterogeneity in studies and risk of bias (134). Certainly, for patients with suspected extraesophageal reflux that fail empiric treatment with PPI therapy, there is a well-accepted role for further testing with pH impedance monitoring because more than 50% of patients may have nonacid reflux (135). At the same time, there is growing interest in examining the utility of up-front diagnostic testing in contrast to empiric PPI trials in this regard. A cost minimization study examining an empiric PPI regimen vs initial physiologic evaluation with pH impedance estimated an average weighted cost of $1,897 for up-front testing and $3,033 for empiric twice daily (BID) PPI and overall a cost-saving with up-front testing (136). In addition to cost-saving, up-front testing may minimize the misdiagnoses of extraesophageal reflux and predict response to fundoplication. In a prospective study of 24 patients with suspected LPR not responsive to PPI therapy, pH impedance findings did not differ compared with controls (137). Another study comparing the diagnostic accuracy of empiric PPI therapy to dual probe pH monitoring for LPR reported a 92.5% sensitivity and 14% specificity of empiric PPI therapy (138). Furthermore, in a retrospective cohort study of 237 patients with extraesophageal symptoms not responsive to PPI therapy, AET on reflux monitoring predicted response to fundoplication (139). These data suggest that patients with suspected LPR not responsive to acid suppression likely do not have LPR pathophysiology, and up-front testing identifies those patients who stand to benefit from antireflux therapy with higher rates of treatment compliance compared with empiric therapy alone (140).

In patients evaluated for chronic cough, observational and outcomes data generally also support up-front testing. A randomized controlled trial of PPI vs placebo of 40 subjects with chronic cough without heartburn found that PPI did not improve chronic cough-related quality of life or symptoms (141). In a prospective study of 30 patients (10 with chronic cough, 10 with GERD, and 10 healthy controls), those responding to PPI were more likely to have weakly acidic esophagopharyngeal reflux and swallowing-induced acidic/weakly acidic esophagopharyngeal reflux (142). Another study of 156 patients with chronic cough undergoing pH impedance found that pathological AET and BI increased the probability of PPI response (143). On the other hand, a study of 27 patients with unexplained chronic cough randomized to high-dose PPI vs placebo found a significant symptom improvement for patients in the PPI arm, regardless of whether they met the criteria for reflux (consisting of endoscopic findings, positive pH impedance study, and/or positive GERDQ) (144). The HASBEER tool reports concomitant asthma, hiatal hernia, heartburn, and rising body mass index as pretest predictors of abnormal pH in patients failing PPI therapy (68). High AET time is uncommon with extraesophageal symptoms, and pH impedance monitoring seems to improve diagnostic yield (145).

Thus, available data suggest that empiric PPI trials may minimally outperform the placebo in patients with suspected extraesophageal reflux. However, pH impedance testing off PPI detects reflux and predicts response to PPI therapy, particularly for patients without typical reflux symptoms. Therefore, up-front ambulatory reflux testing (pH impedance testing off PPI) is a more specific diagnostic approach compared with empiric PPI. Parameters on pH impedance monitoring that have best value for diagnosis of extraesophageal reflux remain unresolved.

**Recommendations**

14. We recommend up-front ambulatory reflux monitoring off acid suppression over an empiric trial of PPI therapy for extraesophageal reflux symptoms without concurrent typical reflux symptoms (conditional recommendation, very low quality of evidence).
Rumination syndrome

Rumination syndrome is diagnosed when patients report repetitive, effortless regurgitation of recently ingested food into the mouth, followed by either spitting or remastication and reswallowing, without nausea, retching, or vomiting (146). The regurgitated food is often recognized and has a pleasant taste. Clinical suspicion and the final diagnosis of rumination syndrome are essentially clinical, using the Rome IV criteria (146). Esophageal function testing (HRIM and pH impedance monitoring) are used to confirm the diagnosis when needed to convince the patients and their caregivers and to rule out confounding diagnoses such as achalasia or primary reflux disease. The sensitivity and specificity of HRIM in the diagnosis of rumination syndrome are 75%–80% and 100%, respectively, based on a study of 15 children and adolescents with rumination syndrome and 15 controls (147,148). The use of a postprandial monitoring protocol increases the likelihood of identification of rumination episodes (91,149). In a retrospective review of 94 patients with persistent esophageal symptoms despite PPI therapy, 20% had a rumination profile during postprandial HRIM monitoring lasting up to 90 minutes after a refluxogenic meal (91). Frequent swallowing during HRIM can confound the diagnosis because of relaxation of the LES (147). The manometric findings consist of an increase in intraesophageal pressure of >30 mm Hg associated with proximal movement of gastric content, esophageal pressurization, and a clinically recognized rumination episode (147,148,150). There is proximal movement of the EGJ from the intra-abdominal cavity into the thorax with the increased intra-abdominal pressure that occurs at the onset of rumination episodes (151).

On pH impedance studies, rumination episodes are not distinguishable from GERD, using standard reflux metrics (150). However, more “reflux” episodes are noted to extend to the proximal esophagus in rumination. Bl values are also similar to GERD and do not provide discrimination (152). In a study of 5 patients with rumination, combined ambulatory high-resolution manometry and pH impedance had 86% sensitivity for identification of rumination episodes, but this technique is not universally available for clinical use (153).

Recommendations

15. We recommend HRIM with postprandial monitoring be used to confirm the diagnosis of rumination syndrome if clinically necessary in patients with esophageal symptoms suspicious for rumination syndrome (conditional recommendation, low quality of evidence).

Supragastric belching

Supragastric belching consists of frequent, repetitive, bothersome belching episodes occurring more than 3 days a week, with the criteria fulfilled for 3 months and symptom onset at least 6 months prior (the Rome IV criteria) (146). pH impedance monitoring is considered the gold standard for the diagnosis of supragastric belching. The presence of air within the esophagus can be identified by rapid increase in intraesophageal impedance, and directionality of air movement is determined by interrogation of data from sequential impedance electrodes (154). HRIM can also identify supragastric belching episodes but only if a postprandial monitoring protocol is used (91). When clinical evaluation was compared with ambulatory pH impedance monitoring, repetitive belching on questioning had a sensitivity of 93.4% and specificity of 75% for a diagnosis of supragastric belching, with a positive predictive value 96.8% and negative predictive value 60.0% (155).

Supragastric belching episodes were identified in 48% of 50 consecutive patients with reflux symptoms at a median rate of 13 episodes/24 hours (interquartile range 6–52) on ambulatory pH impedance monitoring, whereas 50% of 10 normal healthy volunteers had 2 (1–6) episodes (156). When daytime upright and nighttime supine periods on ambulatory pH impedance monitoring were analyzed separately in 14 patients with excessive belching, supragastric belches were identified almost exclusively while upright (37.8 ± 6.1 episodes/hr) compared with supine periods (0.9 ± 0.5 episodes/hr, P < 0.001), demonstrating that supragastric belching is suppressed during sleep (157). By contrast, gastric belches remain constant throughout the 24-hour period. In patients with troublesome belching symptoms, supragastric belches are more frequent than gastric belches (158), and supragastric belches determine the severity of symptoms rather than gastric belches (159). Therefore, identification of supragastric belches on pH impedance monitoring is of value in clinical diagnosis of belching disorders and in planning management.

Recommendations

16. We recommend that for patients with excessive belching, pH impedance monitoring can be used to confirm the diagnosis of supragastric belching (conditional recommendation, very low quality of evidence).

NEW DIAGNOSTIC MODALITIES AND METRICS

There is active investigation to identify diagnostic tools that can reliably identify GERD. Catheter-based oropharyngeal pH monitoring has been proposed as a method to detect supragastric reflex events. The Restech Dx-pH system (Respiratory Technology, San Diego, CA) uses a nasopharyngeal catheter to measure pH in liquid or aerosolized droplets at the posterior oropharynx. In addition to normative data, a composite Ryan score has been developed for this device, consisting of 3 components: the number of reflux episodes, duration of longest reflux episode, and % time spent below a pH threshold of 5.5 in the upright and 5.0 in the supine position (160). However, in both pediatric and adult populations, correlation could not be demonstrated regarding reflux events between esophageal pH impedance monitoring and oropharyngeal pH monitoring (161,162). Differences in oropharyngeal monitoring could not be identified between symptomatic patients and healthy volunteers, and as many as 33% of healthy volunteers had a positive Ryan score (163). Furthermore, oropharyngeal pH monitoring was unable to predict which patients with laryngeal symptoms would respond to acid suppressive therapy (164). These data have tempered the initial enthusiasm for the Restech Dx-pH system as a minimally invasive device for the detection of extraesophageal reflux. Further research in larger well-defined patient populations is needed to better understand the diagnostic utility of oropharyngeal pH monitoring.

Measurement of pepsin concentration in saliva has been proposed as a noninvasive method of detecting gastroesophageal reflux, and particularly extraesophageal reflux. Peptest (RD Biomed, Hull, UK) is a recently marketed diagnostic tool to rapidly quantify salivary pepsin concentrations using a lateral flow device with monoclonal antibodies to human pepsin. An initial assessment of saliva samples collected from 58 patients
with GERD and 51 controls identified acceptable test characteristics with an 81% positive predictive value and 78% negative predictive value (165). In another study with 100 controls and 111 patients with symptomatic heartburn, a saliva sample with salivary pepsin > 210 ng/mL was 98.2% specific for GERD and/or reflux hypersensitivity compared with pH impedance monitoring (166). More recently, 2 studies comparing symptomatic patients with GERD to asymptomatic volunteers identified no significant difference in salivary pepsin concentration between the 2 groups and detected positive salivary pepsin results in most volunteers (167,168). These negative findings have raised concerns related to diagnostic reliability and reproducibility of Peptest. Furthermore, a recent meta-analysis of 11 observational studies assessing Peptest in LPR reported a pooled sensitivity of 64% and specificity of 68%. The meta-analysis was limited by significant heterogeneity across studies (169). Although the precise diagnostic role for salivary pepsin testing remains unclear, salivary pepsin testing has many positive attributes (e.g., noninvasive, rapid, and cost-efficient) and continues to be studied as an alternative diagnostic screening tool for GERD and LPR.

The clinical and investigational value of FLIP continues to expand. However, FLIP studies evaluating EGI barrier function in GERD have not demonstrated a discriminative value for EGI distensibility in segregating symptomatic GERD from controls (170,171). On the other hand, impaired esophageal body contractile response to volumetric distension has been associated with abnormal esophageal acid burden in a small study (172), but more research is needed along similar themes. Intraoperative FLIP during ARS and endoscopic reflux procedures is feasible, and the distensibility index can help tailor the intervention to prevent postoperative reflux (173–177). Thus, FLIP utilization in GERD is in its infancy, and large prospective studies are needed to better define the role of FLIP in GERD management.

Within esophageal reflux testing, novel HRM and pH impedance parameters including EGI contractile integral, provocative testing during HRM, BI, and PSPW introduced in recent years show promise. These need to be studied prospectively to understand if patients stand to benefit from the use of these metrics, especially when conventional metrics are inconclusive. Similarly, evaluation of esophageal MI is interesting as a potential diagnostic tool for GERD. These metrics, especially when conventional metrics are inconclusive, need to better define the role of FLIP in GERD management.

CONCLUSIONS

Esophageal presentations, patient self-report questionnaire information, and even empiric therapeutic trials are not always predictive of esophageal disorders with high certainty. The overall goal of esophageal physiologic testing should be to identify unique characteristics about each symptomatic patient that will allow delivery of precision, personalized management. A major setback is that existing esophageal research evaluating the value of esophageal function tests has low quality of evidence. However, despite low GRADE quality and conditional recommendations, esophageal physiologic testing options form an integral part of patient evaluation in the setting of esophageal symptoms that persist without objective evidence of etiology or pathophysiology on endoscopy. Proper test selection, with understanding of test performance characteristics and limitations, can help identify disease processes and predict symptom outcome from management. Prospective high-quality studies using multiple modalities of esophageal tests for symptomatic patients are needed to understand the true value of each physiologic test in predicting outcome.

ACKNOWLEDGMENT

The authors would like to express gratitude to Carol Shannon, an informationist at the University of Michigan, who performed all the data searches for each of the PICO questions for this guideline and, to the guideline monitor, Sumant Inamdar. This guideline was produced in collaboration with the Practice Parameters Committee of the American College of Gastroenterology. The Committee gives special thanks to Sumant Inamdar, MD, who served as guideline monitor for this document, and to Rena H. Yadlapati, MD, MSHS, who assisted with the GRADE methodology process.

CONFLICTS OF INTEREST

Guarantor of the article: C. Prakash Gyawali, MD.
Specific author contributions: All authors contributed equally to this guideline.
Financial support: None to report.
Potential competing interests: C.P.G.: Medtronic, Diversityspeak (teaching and consulting), Ironwood, Iso-thrive (consulting); D.A.C.: Medtronic (teaching and consulting); also has a licensing agreement with Medtronic; J.C.: A.P.: R.J.W: R.Y.: None to report.

REFERENCES


68. Patel DA, Sharda R, Choksi YA, et al. Model to select on-therapy vs o

73. Savarino E, Zentilin P, Frazzoni M, et al. Characteristics of gastro-

70. Bello B, Zoccali M, Gullo R, et al. Gastroesophageal reflux disease and

64. Patel A, Sayuk GS, Kushnir VM, et al. GERD phenotypes from pH-


77. Bajbouj M, Becker V, Phillip V, et al. High-dose esomeprazole for

60. Spechler SJ, Hunter JG, Jones KM, et al. Randomized trial of medical

62. Patel A, Sayuk GS, Gyawali CP. Parameters on esophageal pH-


58. Shaheen NJ, Denison H, Bjorkc K, et al. Esophageal mucosal breaks in
gastroesophageal reflux disease partially responsive to proton pump

on therapy to select patients with refluxatory symptoms for
antireflux surgery. A single center retrospective study.

patients with reflux symptoms referred for pH and impedance testing

antireflux surgery—what is the proper preoperative work-up?

381:1513–23.


68. Patel DA, Sharda R, Choksi YA, et al. Model to select on-therapy vs o

70. Bello B, Zoccali M, Gullo R, et al. Gastroesophageal reflux disease and
antireflux surgery—what is the proper preoperative work-up?

381:1513–23.


68. Patel DA, Sharda R, Choksi YA, et al. Model to select on-therapy vs o

70. Bello B, Zoccali M, Gullo R, et al. Gastroesophageal reflux disease and
antireflux surgery—what is the proper preoperative work-up?

381:1513–23.


