Ambulatory reflux monitoring for diagnosis of gastro-esophageal reflux disease: Update of the Porto consensus and recommendations from an international consensus group


1Digestive Physiology, Hospices Civils de Lyon and Lyon I University, Inserm U1032, LabTAU, Lyon, France
2Division of Gastroenterology, Washington University School of Medicine, St. Louis, MO, USA
3Division of Gastroenterology, Department of Surgical, Oncological and Gastroenterological Sciences, Padua, Italy
4Division of Gastroenterology, Department of Medicine, Northwestern University, Chicago, IL, USA
5Department of Gastroenterology, Bordeaux University Hospital, and Université de Bordeaux, Bordeaux, France
6Institute of Digestive Disease, The Chinese University of Hong Kong, Hong Kong, China
7Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, AZ, USA
8Division of Gastroenterology, University Clinics for Visceral Surgery and Medicine, Bern University Hospital, Bern, Switzerland
9Department of Surgery, University of Washington, Seattle, WA, USA
10Center for Digestive Diseases, Bart’s and the London School and Dentistry, London, UK
11Department of Internal Medicine, Israeliitic Hospital, University of Hamburg, Hamburg, Germany
12Department of Gastroenterology, Abdominal Center, St. Claraspital, Basel, Switzerland
13Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands

Abstract

Background: An international group of experts evaluated and revised recommendations for ambulatory reflux monitoring for the diagnosis of gastro-esophageal reflux disease (GERD).

Methods: Literature search was focused on indications and technical recommendations for GERD testing and phenotypes definitions. Statements were proposed and discussed during several structured meetings.

Key Results: Reflux testing should be performed after cessation of acid suppressive medication in patients with a low likelihood of GERD. In this setting, testing can be either catheter-based or wireless pH-monitoring or pH-impedance monitoring. In patients with a high probability of GERD (esophagitis grade C and D, histology proven Barrett’s mucosa >1 cm, peptic stricture, previous positive pH monitoring) and persistent symptoms, pH-impedance monitoring should be performed on treatment. Recommendations are provided for data acquisition and analysis. Esophageal acid exposure is considered as pathological if acid exposure time (AET) is greater than 6% on pH testing. Number of reflux episodes and baseline impedance are exploratory metrics

aSee Appendix 1.
Gastro-esophageal reflux disease (GERD) is defined as a “condition which develops when the reflux of gastric content causes troublesome symptoms or complications”. A large proportion of patients are treated empirically with acid suppressant drugs. Further investigation is mainly indicated in patients with warning signs, atypical symptoms, lack of response to pharmacological therapy or prior to invasive endoscopic or surgical procedures. While upper gastrointestinal endoscopy is the first line examination to evaluate reflux consequences (esophagitis, Barrett’s mucosa) and to rule out conditions mimicking reflux symptoms (e.g. tumors, strictures), this cannot measure reflux itself or provide definite evidence that symptoms are caused by reflux; instead, ambulatory reflux monitoring is required for this purpose. The 2004 Porto consensus provided recommendations for reflux definition and detection. Wireless pH monitoring and pH-impedance monitoring were introduced only a few years prior. Since then, monitoring techniques have rapidly gained ground and numerous publications have provided new insights, warranting an update of the Porto consensus.

An international group of experts proposed to revise recommendations for ambulatory reflux monitoring for GERD diagnosis. The conclusions of the working group are presented here.

Key Points

- The Porto consensus proposed recommendations for GERD testing in 2004. An international working group proposed to revise these recommendations.
- Pathological GERD is defined by at least one of the following criteria: Grade C or D esophagitis, peptic stricture, Barrett’s mucosa >1 cm and esophageal acid exposure >6%.
- Number of reflux episodes and baseline impedance should be considered as exploratory tools for further research.

1 | INTRODUCTION

Gastro-esophageal reflux disease (GERD) is defined as a “condition which develops when the reflux of gastric content causes troublesome symptoms or complications”. A large proportion of patients are treated empirically with acid suppressant drugs. Further investigation is mainly indicated in patients with warning signs, atypical symptoms, lack of response to pharmacological therapy or prior to invasive endoscopic or surgical procedures. While upper gastrointestinal endoscopy is the first line examination to evaluate reflux consequences (esophagitis, Barrett’s mucosa) and to rule out conditions mimicking reflux symptoms (e.g. tumors, strictures), this cannot measure reflux itself or provide definite evidence that symptoms are caused by reflux; instead, ambulatory reflux monitoring is required for this purpose. The 2004 Porto consensus provided recommendations for reflux definition and detection. Wireless pH monitoring and pH-impedance monitoring were introduced only a few years prior. Since then, monitoring techniques have rapidly gained ground and numerous publications have provided new insights, warranting an update of the Porto consensus.

An international group of experts proposed to revise recommendations for ambulatory reflux monitoring for GERD diagnosis. The conclusions of the working group are presented here.

2 | METHODS

Following the model of the Chicago Classification for definition of esophageal motility disorders in high resolution manometry (HRM), an international group of experts convened at five distinct meetings (UEGW 2014, Ascona 2015, UEGW 2015, DDW 2016 and UEGW 2016) to establish and discuss guidelines for GERD testing based on extensive literature search. The search was focused on indications for GERD testing, technical recommendations for data acquisition and analysis of esophageal pH and pH-impedance monitoring, and GERD phenotypes. Statements for GERD testing were thus proposed and graded on the quality of the supporting evidence according to the GRADE system (Table 1). The strength of the individual statements is based on the aggregate evidence quality and a balance between benefits and harms. Grade category is indicated in the text within parenthesis in italics and in Tables; details are provided in Data S1.

3 | INDICATIONS FOR GERD TESTING

3.1 | Indications of pH and pH-impedance monitoring

GERD testing is usually performed when there is a need for a definite diagnosis of GERD. It may be indicated in patients with incomplete or lack of response to PPI therapy, prior to anti-reflux surgery, and for atypical symptoms like cough, frequent belching, and suspected rumination (GRADE low).

Recommendations are provided in Table 2 to guide the modality of GERD testing. The use of 24-hours esophageal impedance-pH monitoring is currently considered as the gold standard for the detection of reflux episodes, since impedance measurement permits the detection of anterograde and retrograde bolus (liquid, gas, or mixed) flow in the esophagus and combined-pH monitoring allows the chemical characterization of the refluxate. Therefore, compared to pH-based reflux monitoring alone, pH-impedance monitoring can detect not only acidic (pH<4) but also weakly acidic (4≤pH<7), weakly alkaline (pH≥7), gaseous, and re-reflux episodes. This definitely increases
the diagnostic yield of reflux monitoring in patients with GERD.\textsuperscript{17,18} However availability, cost and patient preference may drive the choice between catheter based pH, pH impedance or wireless pH monitoring.\textsuperscript{19}

Specific indications for wireless pH monitoring include intolerance of the transnasal catheter (GRADE moderate),\textsuperscript{20,21} or a negative catheter based pH study with high suspicion of GERD, to elicit day to day variation in acid exposure and for improving detection of symptom association (GRADE moderate).\textsuperscript{22–24} Prolongation of pH monitoring beyond 24 hours increases sensitivity of reflux detection and symptom events for symptom reflux association.\textsuperscript{25,26}

In case of incomplete or no response to PPI therapy in patients with a high probability of reflux disease, GERD testing is performed on medication. This determines whether the patient is refractory to PPI therapy by quantifying reflux burden on medication, and assesses whether reflux and symptoms are related in time.\textsuperscript{15,27} Thus, using combined impedance-pH monitoring is useful to determine if PPI failure is associated with (i) ongoing acid reflux, (ii) adequate acid control but ongoing symptomatic non-acid reflux, or (iii) no reflux.\textsuperscript{15,27–31}

While pH-monitoring is recommended before anti-reflux surgery,\textsuperscript{7} there are limited data on the incremental value of adding impedance to pH monitoring to facilitate selecting patients for anti-reflux surgery, and even less data on the indication for surgery based on the number of reflux episodes alone.\textsuperscript{32–35} Good surgical outcomes (in terms of patient satisfaction) have been reported when using abnormally high total number of reflux episodes detected by impedance off therapy (in patients with normal pH parameters) as a means of selecting patients for antireflux surgery.\textsuperscript{36} This was also demonstrated in patients

\begin{table}[h]
\centering
\caption{Grading of recommendations assessment, development, and evaluation (GRADE) system for the quality of evidence for guidelines}
\begin{tabular}{|l|l|}
\hline
Quality of evidence & Definition \\
\hline
High quality & Further research is very unlikely to change our confidence in the estimate of effect. \\
\hline
Moderate quality & Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate. \\
\hline
Low quality & Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate. \\
\hline
Very low quality & Any estimate of effect is very uncertain. \\
\hline
\end{tabular}
\end{table}

\begin{table}[h]
\centering
\caption{Recommendations for the choice of GERD testing}
\begin{tabular}{|l|l|}
\hline
Recommendations from the working group & GRADE category \\
\hline
Esophageal pH impedance monitoring may be indicated for interpretation of patients with refractory symptoms despite PPI therapy, prior to and/or following anti-reflux surgery, and for symptoms of cough, frequent belching, and rumination syndrome & Low \\
\hline
pH impedance studies performed on PPI demonstrate a decreased frequency of acid reflux episodes, and an increase in proportion of non-acidic or weakly acidic episodes & Moderate \\
\hline
Data are insufficient to suggest antireflux surgery based solely on an increase in number of reflux episodes detected by impedance & Very low \\
\hline
An absolute indication for wireless pH monitoring is in patients intolerant of a pH or pH impedance catheter & Moderate \\
\hline
Wireless pH monitoring is indicated in patients with a negative catheter based pH study in a patient with ongoing symptoms, to elicit day to day variation in acid exposure and symptom association & Moderate \\
\hline
Reflux monitoring (catheter based pH, wireless pH, or pH impedance) should be performed off of PPI to demonstrate abnormal reflux prior to antireflux surgery & Very low \\
\hline
Reflux monitoring (catheter based pH, wireless pH, or pH impedance) should be performed off of PPI to demonstrate abnormal reflux in the setting of PPI non-response & Very low \\
\hline
Reflux monitoring in the form of pH impedance should be performed on PPI in settings with prior evidence for reflux (prior positive pH testing, esophagitis grade C or D, histology proven Barrett’s esophagus > 1 cm, peptic stricture) & Moderate \\
\hline
Pharyngeal reflux monitoring has no value to guide clinical management & Very low \\
\hline
Manometry with impedance is indicated to distinguish rumination from GERD related regurgitation and to distinguish reflux-cough from cough-reflux sequence & Very low \\
\hline
\end{tabular}
\end{table}
who were studied on PPI. However, these studies lack control groups; there is also no information on whether patients with low total number of reflux episodes perform poorly after antireflux surgery. Furthermore, others found that abnormal pH values and "typical" reflux symptoms were better predictors of a positive outcome after antireflux surgery compared to results of impedance testing off or on medication. Consequently, data are insufficient to recommend antireflux surgery based solely on an increased number of reflux episodes detected by impedance (GRADE very low).

In the evaluation of patients who report recurrent symptoms after surgery, the yield of pH-impedance testing over pH-only monitoring is variable. Some authors suggested that symptoms were driven by weakly acid reflux episodes while others found that reflux episodes detected by impedance monitoring were not significant contributors of symptoms. Interestingly one study demonstrated that supra-gastric belching was the cause of recurrent symptoms after surgery in a third of patients. Moreover, the severity of post-surgery gas-related symptoms was not associated with an increased number of pre-operative air swallows and/or belches or a larger post-operative decrease in the number of gastric belches. In absence of clear data in the literature, the working group recommends the same reflux monitoring modality that was performed prior to surgery be performed following surgery off of PPI (expert recommendation).

Establishing a relationship between chronic cough and reflux is challenging. Chronic cough can be associated with weakly acidic reflux, thus favoring the use of impedance-pH rather than pH alone in this particular setting. A cough detector or ambulatory manometry can be an important tool to distinguish cough-reflux and reflux-cough sequences. However, in the absence of outcome data, it is difficult to determine which patients with chronic cough associated with reflux would benefit from anti-reflux surgery.

Excessive belching is a common phenomenon observed in patients with GERD or functional dyspepsia. Esophageal impedance reliably distinguishes gastric belch episodes from supragastric belch episodes. Interestingly, supragastric rather than gastric belching is frequently encountered in patients with complaint of excessive belching. Thus, impedance-pH monitoring may be considered the "gold standard" for the workup of patients with excessive belching. When performed in conjunction with manometry, impedance can also be used to distinguish rumination syndrome from GERD-related regurgitation in patients with normal lower esophageal sphincter (LES) function who have emesis without antecedent retching as in rumination a gastric pressure increase >30 mm Hg will precede the occurrence of reflux.

3.2 Testing off or on PPI

Determining if GERD testing is to be performed off or on PPI therapy is an important initial decision point in esophageal physiologic testing. Testing off PPI results in more symptoms reported during the measurement and thus a higher yield of symptom-reflux association analysis. Testing off PPI also reveals the naïve esophageal acid exposure time. On the other hand, in some instances one would like to know why the treatment for GERD is not effective and whether complete acid suppression is obtained. The decision should be made with consideration of the patient’s clinical presentation and pre-test evidence for reflux (Figure 1).

Reflux monitoring (catheter based pH, wireless pH, or pH impedance) should be performed off PPI to confirm if the question is whether reflux is the cause of symptoms and if patient truly has GERD in case of non-response to therapy (GRADE very low). This is also the case when testing prior to antireflux surgery (GRADE very low). In the setting of persistent symptoms in patients with non-erosive reflux disease (normal esophageal mucosa), grade A or B esophagitis, Barrett’s esophagus (defined as intestinal metaplasia on esophageal biopsies) segments (<1 cm), atypical presentations, absent or incomplete response to PPI, and recurrent/persistent symptoms after anti-reflux surgery.
In patients with prior evidence of excessive reflux (prior positive pH testing, esophagitis, Barrett’s esophagus >1 cm or peptic stricture) the main question is not whether there is GERD but why the treatment is failing. In this case impedance-pH testing should be performed on PPI in search of ongoing reflux (either acid or non-acid) despite PPI (GRADE moderate).

### 4 | PERFORMING PH AND PH-IMPEDEANCE MONITORING: RECOMMENDATIONS FOR DATA ACQUISITION

Instructions to be provided to patients for data acquisition are summarized in Table 3. All of them are based on expert recommendations in absence of clear evidence in literature.

By convention, the pH or pH impedance catheter is positioned with the distal esophageal pH probe 5 cm above upper border of the manometrically defined lower esophageal sphincter (LES) for catheter based monitoring, as up to 2 cm of movement in either direction can be expected with swallowing and head movement, and this location reduces catheter migration into the stomach. The wireless probe is positioned transorally 6 cm proximal to the squamocolumnar junction identified during endoscopy (or 9 cm above upper border of the manometrically defined LES from the nostrils), which corresponds to the conventional pH positioning 5 cm proximal to the LES.

### 5 | INTERPRETING PH MONITORING

Recommendations for data analysis are reported in Table 4.

#### 5.1 | pH thresholds defining pathological distal esophageal acid exposure

A drop in esophageal pH below 4 is the most discriminative threshold to define a reflux episode. By cumulative summation of time when esophageal pH is below 4, the acid exposure time (AET, % time with pH<4 over the study duration) can be derived. Meal times, fluid ingestions, and artifacts are typically excluded, as ingested acidic material can confound the calculation. A general visual inspection of pH study should be performed to exclude artifacts and to search for catheter displacement or wireless capsule dislodgment, especially in case of high AET.

A composite parameter taking six individual metrics, total AET, upright AET, recumbent AET, number of reflux episodes, reflux episodes with pH<4 for ≥5 minutes, and duration of longest reflux episode, is described as the DeMeester score. Of these, the total AET is the most reproducible and AET overall is more specific compared to other individual components of the DeMeester score. By consensus, the AET is favored as the metric used to designate esophageal acid burden, and elevated AET has been demonstrated to predict a positive response to PPI trial and symptom outcome following antireflux therapy.

On catheter based pH monitoring off antisecretory therapy, mean total AET values and 95th percentile of normal have ranged from 3.9% to 5.9%.

### TABLE 3

<table>
<thead>
<tr>
<th>Recommendations from the working group</th>
<th>GRADE category</th>
</tr>
</thead>
<tbody>
<tr>
<td>For ambulatory reflux monitoring, the minimum duration of recording for adequate impressions should be 16 hours or more</td>
<td>Low</td>
</tr>
<tr>
<td>In ambulatory reflux monitoring, upright and recumbent periods should be reported separately, with exclusion of mealtimes</td>
<td>Moderate</td>
</tr>
<tr>
<td>Automated analysis of pH impedance studies is adequate for acid reflux events</td>
<td>Very low</td>
</tr>
<tr>
<td>Automated analysis of pH impedance overestimates non-acidic or weakly acidic events</td>
<td>Very low</td>
</tr>
<tr>
<td>Manual review of the 2 minutes preceding each symptom event in pH impedance studies is necessary</td>
<td>Very low</td>
</tr>
<tr>
<td>Low baseline impedance makes interpretation of pH-impedance studies difficult</td>
<td>Low</td>
</tr>
<tr>
<td>Baseline impedance &lt;500 ohms might suggest an additional process, e.g. Barrett’s esophagus, a motor disorder, eosinophilic esophagitis, inflammation, fibrosis, etc.</td>
<td>High</td>
</tr>
</tbody>
</table>

### TABLE 4

<table>
<thead>
<tr>
<th>Recommendations from the working group</th>
<th>GRADE category</th>
</tr>
</thead>
<tbody>
<tr>
<td>For ambulatory reflux monitoring, the minimum duration of recording for adequate impressions should be 16 hours or more</td>
<td>Low</td>
</tr>
<tr>
<td>In ambulatory reflux monitoring, upright and recumbent periods should be reported separately, with exclusion of mealtimes</td>
<td>Moderate</td>
</tr>
<tr>
<td>Automated analysis of pH impedance studies is adequate for acid reflux events</td>
<td>Very low</td>
</tr>
<tr>
<td>Automated analysis of pH impedance overestimates non-acidic or weakly acidic events</td>
<td>Very low</td>
</tr>
<tr>
<td>Manual review of the 2 minutes preceding each symptom event in pH impedance studies is necessary</td>
<td>Very low</td>
</tr>
<tr>
<td>Low baseline impedance makes interpretation of pH-impedance studies difficult</td>
<td>Low</td>
</tr>
<tr>
<td>Baseline impedance &lt;500 ohms might suggest an additional process, e.g. Barrett’s esophagus, a motor disorder, eosinophilic esophagitis, inflammation, fibrosis, etc.</td>
<td>High</td>
</tr>
</tbody>
</table>
to 7.2% in patients who are either asymptomatic or with occasional (<2 episodes/month) heartburn symptoms, with sensitivity of 77%-100% and specificity of 85%-100% in discrimination of esophagitis from normal controls. Based on these data, a total AET value of <4% is consistently normal (GRADE moderate) and less than 6% is likely to be normal (Figure 2). In contrast, patients with erosive esophagitis have mean total AET of 9.6%-27.6%,55,58,61,62,64,66 Thus, the consensus group concluded that AET >6% was consistently abnormal (GRADE high). There is consistent overlap between normal controls and symptomatic GERD without esophagitis within a gray area consisting of 95th percentile AET values of 4%-6%,57,62,68,69 where additional evidence from alternate testing may provide further confidence in the presence of pathologic acid burden (GRADE moderate). Furthermore, there is considerable day-to-day variability in AET measurements so a clinical decision should never be made exclusively based on this parameter.70,71

Wireless pH monitoring is associated with marginally higher 95th percentile AET values in normal controls (4.4%-5.3%),22,72,73 but within the gray area indicated above. Therefore, similar concepts can be applied to distal AET thresholds with wireless pH monitoring as for catheter based pH monitoring. Using wireless pH-monitoring, analysis could be performed at every day separately taking into account the worst day or averaging the entire study period. While sensitivity is obviously higher with worst day analysis, specificity is obviously higher with averaged AET.26,71

5.2 | Upright and supine acid exposure time

Abnormal supine AET is observed in different conditions: poor sleep quality, severe erosive esophagitis, Barrett’s esophagus,74-76 elevated BMI,77 and consumption of a late-evening meal.78 Pre-operative pathological bi-positional (upright and supine) AET may have higher likelihood of recurrent pathological acid exposure and esophagitis following Nissen fundoplication and the need of re-operation, but the results are conflicting.79-81

While upright and supine AET might differ according to associated conditions, further studies are needed to evaluate the value of additional reporting of acid exposure time during sleep.

5.3 | Proximal acid esophageal exposure

Dual sensor pH monitoring was proposed to evaluate proximal and distal reflux burden in patients with extra-esophageal reflux symptoms in particular. Proximal AET has poor sensitivity and reproducibility and cannot predict symptoms severity.83 Furthermore, there is no consensus on pH criteria for defining pathologic reflux at the proximal esophagus.84,85 Thus, there is no clear evidence that dual probe pH monitoring is of additional value above distal pH probe measurement alone.

6 | INTERPRETING PH-IMPEdANCE MONITORing

6.1 | Thresholds to define pathological pH-impedance monitoring

Standardized interpretation of impedance-pH measurement data are based on published normal values.67-69,86 Similar AET thresholds are used for impedance-pH monitoring as for pH monitoring alone (GRADE low) and the same thresholds are used when impedance-pH monitoring is performed off or on therapy (GRADE low)86,87 (Figure 2).

Published thresholds for abnormal number of reflux episodes during impedance-pH monitoring are quite heterogeneous and outcome data providing their usefulness in GERD management are scant.88 Given the fairly good reproducibility of this impedance parameter and the preliminary data correlating symptom remission with the decreasing number of reflux episodes in the post-surgical setting, a clearly high number of reflux episodes (above 80) might be considered abnormal (GRADE low), while a number of reflux episodes on impedance-pH of 40 or fewer are considered as normal (GRADE low).
low). However, number of reflux episodes alone is not predictive of treatment outcome.\textsuperscript{33,38,40} Therefore, caution should be adopted to diagnose GERD based on numbers of reflux events alone, and additional clinical and investigation findings should be considered (GRADE low).\textsuperscript{90} Consequently, the consensus group recommends reporting number of reflux episodes detected on impedance as an adjunctive tool rather than a primary indicator of abnormal reflux burden.

### 6.2 | Analysis of pH-impedance monitoring

Recommendations are described in Table 4. After completion of the impedance-pH study, data are analyzed using proprietary software and interpreted by the reporting physician. The software identifies individual reflux and swallow events, measures symptom-reflux association, and distinguishes changes in impedance that are not clinically important.\textsuperscript{95} Automated analysis is adequate for acid reflux events, but overestimates non-acidic or weakly acidic events. As a consequence, calculation of Symptom Index (SI) and Symptom Association Probability (SAP) might be affected.\textsuperscript{92,93} A manual review of the 2 minutes preceding each symptom event in pH impedance studies may classify most of the patients similarly as manual analysis of the 24-hours study regarding the positivity of SI and SAP.\textsuperscript{94} Therefore, manual editing of the 2 minutes preceding each symptom event is recommended (GRADE very low).

### 6.3 | Baseline impedance

Low baseline esophageal impedance is observed in case of impaired esophageal mucosal integrity (erosive esophagitis, Barrett’s esophagus, eosinophilic esophagitis) and in case of bolus stasis secondary to severe esophageal motility disorders such as achalasia, absent peristalsis, severe ineffective esophageal motility.\textsuperscript{2,16,95,96} Low baseline impedance value (<500 ohms) may affect the sensitivity of reflux recognition during pH-impedance monitoring. While it does not preclude the recognition of reflux during visual analysis, it can make automated recognition difficult (GRADE low) and a manual analysis with magnification of the tracings may help in reflux episodes identification. Furthermore, as low baseline impedance might suggest an additional esophageal process and/or disease, it should be reported in impedance studies.

Recently, different studies suggested that measurement of mucosal esophageal impedance could be useful in the evaluation of patients with suspected GERD. Patients with erosive and non-erosive GERD have lower average baseline impedance values than healthy subjects and patient with functional heartburn.\textsuperscript{97–99} Some authors proposed a simplified method to measure baseline impedance. This measurement called mean nocturnal baseline impedance (MNBI) consist of measuring baseline impedance 3 or 5 cm above the LES during the overnight rest as the mean baseline impedance of three 10-minute time periods in a period without swallowing.\textsuperscript{100} MNBI was lower in suspected GERD patients with typical symptoms who responded to PPI therapy or anti-reflux surgery compared to those who did not respond.\textsuperscript{100–104} Interestingly, the threshold values for abnormal impedance baseline values (<2100 ohms) were quite similar to those obtained by means of an impedance catheter introduced in the working channel of an endoscope.\textsuperscript{105} Finally, recent studies found that impedance changes due to acute esophageal inflammation might be reversed with mucosal healing or gastro-esophageal reflux inhibition supporting a role of mucosal impedance as a hallmark of pathological reflux disease.\textsuperscript{106,107}

Since baseline impedance is deemed to have potential as a metric for predicting outcome (GRADE moderate), the consensus recommends reporting of MNBI as an exploratory tool, and evaluating its yield over AET and symptom association in large prospective studies with outcome data.

### 6.4 | Post-reflux swallow induced peristaltic wave

Evaluating the efficacy of esophageal chemical clearance might be considered to further characterize GERD. Using 24 hours pH-impedance monitoring, post-reflux swallow induced peristaltic wave (PSPW) index was proposed to evaluate chemical clearance. Briefly PSPW is defined as an antegrade 50% drop in impedance relative to the pre-swallow baseline originating in the most proximal impedance sites, reaching all the distal impedance sites, and followed by at least 50% return to the baseline in all the distal impedance sites (bolus exit).\textsuperscript{108} Further PSPW index is calculated manually as the number of refuxes followed within 30 seconds by a PSPW divided by the number of total reflux. PSPW index is significantly lower in patients with reflux esophagitis and NERD compared to controls and patients with functional heartburn.\textsuperscript{109,110} While PSPW index may have complementary value, particularly when pH-impedance is performed on PPI (GRADE low),\textsuperscript{104} there is currently not enough evidence for its clinical use.

### 7 | PHARYNGEAL PH-(IMPEDE) MONITORING

The role of pharyngeal reflux monitoring for the diagnosis of gastro-esophageal reflux episodes extending to the pharynx remains unclear.\textsuperscript{46} There is no consensus regarding definition of pharyngeal reflux\textsuperscript{111} and important limitations (inaccurate catheter position, artifacts, poor reproducibility).\textsuperscript{86,112–117} Measuring aerosolized pharyngeal acid reflux was proposed but considerable disagreement in the detection of reflux between pharyngeal acid aerosol monitoring and impedance was observed.\textsuperscript{118,119} Thus, measurement of airway and pharyngeal pH cannot be recommended to diagnose gastro-esophageal reflux episodes extending to the pharynx (GRADE moderate).\textsuperscript{120,121}

### 8 | SYMPTOM REFLEX ASSOCIATION

Symptom reporting during ambulatory 24-hours reflux monitoring allows investigation of the temporal relationship between reflux and symptom. In this consensus document, the expert group agreed to use the terms symptom events and reflux episodes. Recommendations are presented in Table 5.
Only symptoms that can reasonably be related to reflux episodes such as cough, chest pain, heartburn, and regurgitation are considered for symptom reflux association analysis, whereas symptoms that are definitely not reflux related are not considered (e.g., headache). It is also not possible to perform reliable symptom reflux association analysis for symptoms that lack a crisp onset and are chronically present, such as dyspnea or hoarseness. As far as possible, the dominant symptom or the most bothersome symptom should be used for symptom reflux association (expert recommendation). When the dominant symptom is cough, adding a cough detector (for example catheter with ambulatory pressure sensors to the pH-impedance catheter) can be helpful to identify the exact timing of the cough events and makes it possible to distinguish a cough-reflux sequence from reflux-cough (GRADE very low).\(^4\)

The outcome of symptoms reflux association is more reliable when the patient has reported multiple symptom events (at least three per symptom should be reported for a reliable reflux-symptom association analysis). Therefore, patients should be encouraged to induce many symptoms and report them all. Measurement off PPI usually results in a higher number of symptoms reported as well, and therefore provide a higher chance of performing adequate symptom reflux association analysis.\(^{122}\) There is no upper threshold for number of symptoms allowed. Symptom reflux association is reported for the entire duration of the pH or pH-impedance study, and is not broken down by upright or recumbent periods (GRADE moderate).

When pH testing is used, only pH drops below the threshold of pH 4 are used to designate reflux episodes, and drops of 1 pH unit not reaching the threshold of pH 4 do not constitute a reflux episode (GRADE moderate). When pH impedance testing is used for symptom reflux association, all reflux events detected by impedance is used in calculation of reflux episodes (GRADE high).

<table>
<thead>
<tr>
<th>Recommendations from the working group</th>
<th>GRADE category</th>
</tr>
</thead>
<tbody>
<tr>
<td>The most bothersome or dominant symptom being studied should be used for symptom reflux association</td>
<td>Consensus/Evidence lacking</td>
</tr>
<tr>
<td>Separate symptom reflux association testing for a secondary symptom is possible, but not for more than 2 symptoms</td>
<td>Consensus/Evidence lacking</td>
</tr>
<tr>
<td>At least 3 events per symptom must be reported for calculation of symptom-reflux association</td>
<td>Consensus/Evidence lacking</td>
</tr>
<tr>
<td>Symptom reflux association is reported for the entire duration of the pH or pH impedance study, and is not broken down by upright or recumbent periods</td>
<td>Moderate</td>
</tr>
<tr>
<td>The only time window to be used for symptoms following a reflux event is 2 minutes</td>
<td>Moderate</td>
</tr>
<tr>
<td>When pH testing is used, only pH drops below the threshold of pH 4 are used to designate reflux episodes, and drops of 1 pH unit not reaching the threshold of pH 4 do not constitute a reflux episode</td>
<td>Moderate</td>
</tr>
<tr>
<td>When pH impedance testing is used for symptom reflux association, all reflux events detected by impedance is used in calculation of reflux episodes</td>
<td>High</td>
</tr>
<tr>
<td>Symptom index (SI) and Symptom Association Probability (SAP) have value in pH and pH-impedance monitoring</td>
<td>High</td>
</tr>
<tr>
<td>SI and SAP are complementary and cannot be directly compared to each other</td>
<td>Very low</td>
</tr>
<tr>
<td>For all reflux monitoring, the 2 minute period prior to each symptom event and 2 minute period following each reflux episode should be evaluated prior to calculating the SI</td>
<td>Moderate</td>
</tr>
<tr>
<td>The Ghillebert Probability Estimate (GPE) can substitute the Weusten method of calculation of SAP when necessary</td>
<td>Very low</td>
</tr>
<tr>
<td>Abnormal AET with both SAP and SI positive represents the strongest evidence for reflux</td>
<td>Moderate</td>
</tr>
<tr>
<td>SAP and SI both positive represents stronger symptom reflux association compared to either alone</td>
<td>Very low</td>
</tr>
<tr>
<td>Evaluate SI only if SAP is positive</td>
<td>Low</td>
</tr>
<tr>
<td>There is weak predictability of PPI response with a positive symptom reflux association parameter, particularly SAP</td>
<td>Low</td>
</tr>
<tr>
<td>If SAP and SI are both positive, the probability of PPI response is greater than if both tests are negative</td>
<td>Low</td>
</tr>
</tbody>
</table>

**TABLE 5** Recommendations to assess reflux-symptom association
events detected by impedance are used in calculation of reflux episodes (GRADE high).

In case of few symptom events and few reflux episodes, analysis can be done by quick overview, but in the majority of cases, a quantitative expression of the strength of the relation between symptom events and reflux episodes is needed.\textsuperscript{123} This can be done using the symptom-reflux association parameters, Symptom Index (SI) and Symptom Association Probability (SAP). The Symptom Sensitivity Index (SSI) has limited added value to SI and SAP and is not further discussed.\textsuperscript{124–126} The Binomial Symptom Index (BSI) (also named Ghillebert Probability Estimate, GPE) is a statistical formula that can express the probability that symptom events and reflux episodes are related.\textsuperscript{127} The BSI and SAP are very strongly related and both could be used interchangeably.\textsuperscript{128}

The SI is defined as the percentage of symptom events that are related to reflux episodes, thus number of reflux related symptom events divided by total number of symptom events times 100%. The most often used cut-off is 50%, which means that above 50% the SI is considered positive, i.e. a large proportion of the patient's symptoms are considered to be reflux-related.

The SAP is a statistical parameter that expresses the strength of the relationship between symptom events and reflux episodes during the measurement.\textsuperscript{125} The calculation is more complex than the SI and cannot be done manually but is calculated instead by the measurement software. The cut-off for the SAP is 95%, and a SAP above 95% (corresponding to $P<0.05$, applying Fisher’s exact test on a $2 \times 2$ table) is considered positive for a relationship between symptom events and reflux episodes.

SI and SAP have a predictive value for the effect of medical and surgical treatment of reflux disease, and this is independent of acid exposure time (GRADE high).\textsuperscript{40,129,130} The result of symptom reflux association analysis has a high degree of reproducibility when repeated, at least as high as the result of the acid exposure time.\textsuperscript{131} These indices have some limitations, especially related to day-to-day variability of reflux burden and occurrence of symptom events during the monitoring day.\textsuperscript{132} Further only a minority of symptoms is associated with reflux episodes.\textsuperscript{133} The SI and SAP are complementary and cannot be directly compared to each other (GRADE very low) as they measure different things. The presence of positive SI and positive SAP together provides the best evidence of a clinically relevant association between reflux events and symptoms (GRADE very low). If one test is positive and the other is negative, this represents a grey area and further interpretation with other parameters (AET, number of reflux episodes, baseline impedance...) and clinical factors are necessary.

9 | GERD DIAGNOSIS: DEFINING DIFFERENT PHENOTYPES

Three key questions require consideration when defining GERD phenotypes: (i) Is there evidence of GERD as indicated by esophageal mucosal lesion or increased reflux burden; (ii) Is there an association between reflux episodes and symptom events; and, (iii) Are there non-GERD disorders, (behavioral or functional disorders) that could explain the symptoms?

9.1 | Esophageal mucosal lesion and mucosal integrity

Los Angeles grade C or D esophagitis provides robust evidence of GERD,\textsuperscript{134} and also predicts the response to treatment.\textsuperscript{135–137} Grade A alone is not sufficient for the diagnosis of GERD as it may be encountered in asymptomatic controls.\textsuperscript{138} While anti-secretory therapy is frequently prescribed to patients with Grade B esophagitis, caution is advised in referring patients to surgery based solely on Grade B esophagitis. Indeed grade B esophagitis might be encountered in asymptomatic controls. Further progression from grade A/B esophagitis to grade C/D and Barrett’s esophagus is observed in 1% to 6% and 1% to 12%, respectively, while it may regress to normal endoscopy in 20% to 60%.\textsuperscript{139} Thus, grade B esophagitis is not sufficient to unambiguously proof that a subject symptoms are caused by GERD. Other endoscopic findings such as peptic stricture and Barrett’s esophagus $>1$ cm are indicative of pathological GERD.\textsuperscript{140}

Most of patients with GERD symptoms have normal appearing mucosa on endoscopy, and examination of biopsies through conventional histology has shown limited value (i.e., low specificity) in documenting GERD.\textsuperscript{141,142} The latest Rome classification stated that the exclusion of eosinophilic esophagitis (by performing esophageal biopsies) was required to diagnose esophageal functional disorders to perform esophageal biopsies to exclude eosinophilic esophagitis.\textsuperscript{143} However, there is limited data regarding the clinical utility of this approach as the majority of patients with eosinophilic esophagitis had endoscopic abnormalities and/or dysphagia.

9.2 | Measures of reflux burden

A challenging scenario is encountered when patients have symptoms suggestive of GERD, but normal endoscopy, and incomplete response to acid suppression. In these patients, ambulatory reflux monitoring is useful to define GERD phenotypes and to guide treatment, particularly in patients with PPI-refractory symptoms.\textsuperscript{29,143}

In this consensus, we consider that AET greater than 6% during ambulatory pH-(impedance) monitoring establishes the diagnosis of GERD. The total number of reflux episodes alone and baseline impedance is not sufficient to confirm the diagnosis of GERD and it should be considered as an exploratory tool.

9.3 | Measures of association between reflux and symptoms

A positive SI and/or SAP support an association between reflux episodes and symptom events, which is an independent predictor of response to treatment in retrospective, uncontrolled, studies. In patients without evidence of pathological GERD (i.e., normal endoscopy, normal AET, normal number of reflux episodes), a positive symptom
association study for heartburn or regurgitation suggests reflux hypersensitivity as defined by Rome IV criteria. 

9.4 | Non-GERD functional and behavioral disorders

In some patients with symptoms but completely normal evaluation, symptoms may be due to a functional gastro-intestinal disorder such as functional heartburn or functional chest pain. Behavioral disorders such as rumination or supragastric belching might also explain symptoms. While these disorders are not part of the GERD spectrum, they might be encountered in patients with symptoms suggestive of GERD, especially those who are refractory to PPI. An overlap might also exist between GERD and a functional or behavioral disorder.

9.5 | Algorithmic approach to GERD phenotypes

GERD definition is based on presence of GERD symptoms, endoscopic findings and results of pH-(impedance) monitoring. Figure 3 summarizes GERD phenotypes. However, there are some borderline cases in which the diagnosis of GERD remains uncertain (grade A and B esophagitis, AET between 4% and 6%). Adding testing such as total number of reflux episodes, baseline impedance, histological evaluation, or evaluation of microscopic esophagitis might be considered to determine if there is other argument in favor of GERD in these patients. An alternative approach to phenotype GERD might be to consider GERD mechanisms and high resolution manometry metrics as presented in the accompanying article on motor findings in GERD (Gyawali et al, unpublished data). Finally, where there is a strong suspicion of GERD, repeating a study could also be considered as a negative result might be secondary to intermittent symptoms and day to day variability.

10 | CONCLUSION

Reflux monitoring can have a decisive role in the management of patients with symptoms suggestive of GERD. In combination with upper GI endoscopy, it is useful to establish the diagnosis of GERD and to define different phenotypes that may guide the treatment. The working group proposes a definition of pathological GERD. New concepts here are the introduction of areas of uncertain diagnosis (grade A and B esophagitis and AET between 4% and 6%) and exploratory tools for research (total number of reflux episodes, mean nocturnal baseline impedance). Based on GRADE category, the level of evidence is usually very low to moderate and most recommendations rely on expert consensus. Future research might be focused on outcome data to determine if GERD phenotypes are useful in the management of patients with GERD symptoms.
ACKNOWLEDGMENTS

Guarantor of the manuscript: Sabine Roman. The consensus process was supported by the Working Group for Gastrointestinal Motility and Function. This group received an educational grant of United European Gastroenterology (UEG).

DISCLOSURE

SR has served as consultant for Sandhill Scientific and Medtronic. CPG has served as a consultant for Medtronic and Torax, and has received research funding from Medtronic. ES has served as consultant for Medtronic, Abbvie, Takeda, Otsuka, Malesci. RW none. FZ Consultant Allergan, Reckitt Benckiser, Research funding: Medtronic, Sandhill Scientific. JW none to declare. MV Consultant: Medtronic, Torax, Sandhill Scientific. RTu none to declare. RTa none. DS has received research grants from Sandhill Scientific and Reckitt Benckiser. JK has served as a consultant for Medtronic and for Standard Instruments, and has received research funding. MF Consultant and speaker: Given Imaging, Reckitt Benckiser and Shire Pharmaceuticals. Research funding: Given Imaging, Mui Scientific, Reckitt Benckiser, Astra Zeneca and Nestle. Educational activities: Given Imaging, Sandhill Scientific Instruments and Medical Measurement Systems. JEP has as consultant and speaker for Medtronic and Sandhill, as speaker for Takeda and Astra Zeneca and as consultant for Ironwood and Impleo. AJB has served as a consultant for Medical Measurement Systems and has received research funding from Medtronic and Medical Measurement Systems.

AUTHOR CONTRIBUTION

SR, CPG, ES, JEP, and AJB study design, manuscript drafting, literature search, critical review of content, approval of final manuscript; RW, FZ, JW, MV, RTu, RTa, DS, JK, and MF literature search, critical review of content, approval of final manuscript; GERD consensus group: critical review of content, approval of final manuscript.

REFERENCES


75. Frazzoni M, De Micheli E, Savarino V. Different patterns of oesophageal acid exposure distinguish complicated reflux disease from either erosive reflux oesophagitis or non-erosive reflux disease. Aliment Pharmacol Ther. 2003;18:1091-1098.


SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Roman S, Gyawali CP, Savarino E, et al. Ambulatory reflux monitoring for diagnosis of gastroesophageal reflux disease: Update of the Porto consensus and recommendations from an international consensus group. Neurogastroenterol Motil. 2017;00:e13067. https://doi.org/10.1111/nmo.13067

APPENDIX 1

GERD consensus group