January 9, 2018

Recent Advances in Clinical Practice

Modern Diagnosis of GERD: The Lyon Consensus

C Prakash Gyawali¹, Peter Kahrilas², Edoardo Savarino³, Frank Zerbib⁴, Francois Mion⁵, André Smout⁶, Michael Vaezi⁷, Daniel Sifrim⁸, Mark Fox⁹, Marcelo F. Vela¹⁰, Radu Tutuian¹¹, Jan Tack¹², Albert J Bredenoord⁶, John Pandolfino², Sabine Roman⁵

¹Division of Gastroenterology, Washington University School of Medicine, St. Louis, MO, USA
²Division of Gastroenterology, Department of Medicine, Northwestern University, Chicago, IL, USA
³Division of Gastroenterology, Department of Surgical, Oncological and Gastroenterological Sciences, Padua, Italy
⁴Department of Gastroenterology, Bordeaux University Hospital, and Université de Bordeaux, Bordeaux, France
⁵Université de Lyon, Hospices Civils de Lyon, Digestive Physiology, Hôpital E Herriot, F-69437 LYON, France; Université de Lyon, Lyon I University, Digestive Physiology, F-69008 LYON, France; Université de Lyon, Inserm U1032, LabTAU, F-69008 LYON, France
⁶Department of Gastroenterology and Hepatology, Academic Medical Center, Amsterdam, The Netherlands
⁷Division of Gastroenterology, Vanderbilt University, Nashville, TN, USA
⁸Barts and The London School of Medicine and Dentistry Queen Mary, University of London, UK
⁹Gastroenterology, St. Claraspital, Kleinriehenstrasse 30, Basel, Switzerland
¹⁰Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, AZ, USA
¹¹Division of Gastroenterology, University Clinics for Visceral Surgery and Medicine, Bern University Hospital, Bern, Switzerland
¹²Department of Gastroenterology, Catholic University of Leuven, Leuven, Belgium

Address all correspondence to C. Prakash Gyawali, MD, MRCP
Professor of Medicine
Washington University School of Medicine
660 S. Euclid Ave, Campus Box 8124
St. Louis, MO 63110
Email: cprakash@wustl.edu
Phone: 314 454 8201
Fax: 314 362 3643

Contributorship: All authors contributed to the content of the manuscript, and reviewed, edited and approved the final draft.
DISCLOSURES

**C. Prakash Gyawali:** Consulting: Ironwood, Torax, Quintiles; Teaching and speaking: Medtronic, Diversatek, Reckitt-Benckiser

**Edoardo Savarino:** Consulting: Abbvie, Allergan, MSD, Takeda, Sofar, Janssen; Teaching and speaking: Medtronic, Reckitt-Benckiser, Malesci, Zambon

**Frank Zerbib:** Research support: Medtronic, Sandhill Scientific; Consulting: Allergan, Reckitt-Benckiser; Speaking and teaching: Ipsen Pharma, Biocodex, Coloplast, Takeda, Vifor Pharma, Mayoly Spindler

**Peter Kahrilas:** Consulting: Ironwood

**Francois Mion:** Teaching & speaking: Laborie, Medtronic; Consulting: Allergan, Endostim

**André Smout:** none

**Michael Vaezi:** Vanderbilt University and Diversatek co-own patent on mucosal impedance technology

**Daniel Sifrim:** Research support: Diversatek, Reckitt-Benckiser; OMOM®, Jinshan Science & Technology (Group) Co. Ltd., Chongqing, China


**Marcelo Vela:** consulting: Torax

**Radu Tutuian:** Teaching: Laborie

**Jan Tack:** Consulting: Ironwood

**Arjan Bredenoord:** Research support: Danone, Bayer; speaking and/or consulting: MMS, Astellas, AstraZeneca, Bayer, Almirall and Allergan

**John Pandolfino:** Research support: Impleo; speaking and/or consulting: Medtronic, Diversatek, Torax, Ironwood, Takeda, Astra Zeneca; stock options: Crospon

**Sabine Roman:** Research support: Sandhill Scientific, Crospon; Teaching: Medtronic; Speaker: Mayoly Spindler
ABSTRACT

Clinical history, questionnaire data and response to antisecretory therapy are insufficient to make a conclusive diagnosis of GERD in isolation, but are of value in determining need for further investigation. Conclusive evidence for reflux on esophageal testing include advanced grade erosive esophagitis (LA grades C and D), long-segment Barrett’s mucosa or peptic strictures on endoscopy, or distal esophageal acid exposure time (AET) >6% on ambulatory pH or pH-impedance monitoring. A normal endoscopy does not exclude GERD, but provides supportive evidence refuting GERD in conjunction with distal AET <4% and <40 reflux episodes on pH-impedance monitoring off PPI. Reflux-symptom association on ambulatory reflux monitoring provides supportive evidence for reflux triggered symptoms, and may predict a better treatment outcome when present. When endoscopy and pH- or pH-impedance monitoring are inconclusive, adjunctive evidence from biopsy findings (histopathology scores, dilated intercellular spaces), motor evaluation (hypotensive lower esophageal sphincter, hiatus hernia and esophageal body hypomotility on high-resolution manometry) and novel impedance metrics (baseline impedance, post-reflux swallow induced peristaltic wave index) can add confidence for a GERD diagnosis; however, diagnosis cannot be based on these findings alone. An assessment of anatomy, motor function, reflux burden and symptomatic phenotype will therefore help direct management. Future GERD management strategies should focus on defining individual patient phenotypes based on the level of refluxate exposure, mechanism of reflux, efficacy of clearance, underlying anatomy of the esophagogastric junction, and psychometrics defining symptomatic presentations.

Keywords: gastroesophageal reflux disease; pH-impedance monitoring; manometry; endoscopy
INTRODUCTION

Gastroesophageal reflux disease (GERD) has an estimated worldwide prevalence of 8-33%, involves all age groups and both genders\(^1\), and carries a price tag estimated at >$9-10 billion/year in the US alone, largely related to proton pump inhibitors (PPI) use and diagnostic testing\(^2\). The current paradigm of GERD diagnosis hinges on the identification of esophageal mucosal lesions or troublesome symptoms caused by gastroesophageal reflux\(^3\). A putative GERD diagnosis is bolstered by a favorable response to PPI therapy\(^4\). The primary determinant of mucosal injury is excessive esophageal acid exposure attributable to anatomical or physiological defects of the esophagogastric junction (EGJ) and esophageal peristalsis\(^5\). GERD symptoms, however, have multiple potential determinants including the number of reflux episodes, the proximal extent to which the refluxate migrates, the acidity of the refluxate, esophageal hypersensitivity, and cognitive hypervigilance. Consequently, depending on the clinical context, the defining features of GERD can be pathology, physiology, or symptomatology. In this paradigm, esophageal testing is often undertaken to define optimal management, be that PPI therapy, antireflux surgery (ARS), or cognitive behavioral therapy.

The aim of the GERD consensus process was to determine modern indications for esophageal testing in GERD, and as an extension to that aim, to define criteria for a clinical diagnosis of GERD. The consensus process started in 2014, when the primary aims were formulated, followed by a literature search and grading of evidence. There was extensive discourse within a multi-national group of GERD experts over 2 years, following which consensus statements were developed and published\(^6^8\). These consensus statements were adapted for the practicing gastroenterologist by a cohort of international experts at the Lyon GERD consensus meeting in November 2017, the conclusions from which are presented in this manuscript.
DIAGNOSIS OF GERD

GERD is empirically diagnosed and treated in clinical practice based on the clinician’s symptom assessment. Indications for testing include treatment failure, diagnostic uncertainty, and treating (or preventing) complications of GERD. However, diagnostic testing may or may not support the initial diagnosis, as the criteria defining GERD are specific to each testing modality. Consequently, understanding the performance characteristics of each diagnostic modality and recognizing evidence that supports or refutes the clinical impression of GERD is crucial. The Lyon Consensus evaluated GERD diagnostic tests from that perspective, and test results were categorized as being adequate to establish or refute a GERD diagnosis or inconclusive in the absence of additional supportive evidence.

The primary focus of esophageal testing has hitherto been restricted to detection of excessive acid reflux as indicative of pathologic GERD, supported by reflux-symptom association analysis. The threshold value discriminating abnormal from normal esophageal acid exposure with 24-hour pH-metry was initially selected based on evidence that higher acid exposure is associated with the presence of reflux esophagitis\(^9\). However, the focus of current GERD management is not only on healing mucosal disease but also on managing symptoms. The association of esophageal acid exposure with patient symptoms is weak\(^10,11\), making sole reliance of this metric problematic. Expansion of testing to include esophageal mucosal impedance, manometry, histopathology, and psychometrics may help in this regard, potentially identifying distinct GERD phenotypes with unique management implications. Each test adds a piece to the overall puzzle of symptom generation, disease pathophysiology, and precision management.

Clinical history and questionnaires

Typical GERD symptoms (heartburn and acid regurgitation) are more likely than atypical symptoms to respond to treatment, emphasizing the value of an accurate clinical history\(^5\). However, when compared to objective evidence of GERD defined by pH-metry or endoscopy, even an expert history by a
gastroenterologist has only 70% sensitivity and 67% specificity\textsuperscript{12}, reiterating the distinction between a physiology-based and a symptom-based GERD diagnosis. Likewise, questionnaires such as the reflux disease questionnaire (RDQ) and GERDQ have similar limitations when compared to physiologic testing\textsuperscript{12-14}. However, in clinical practice, diagnosing and treating GERD based on typical symptoms is pragmatic and endorsed by societal guidelines\textsuperscript{4}, even though these symptoms are neither sensitive nor specific for objectively defined GERD\textsuperscript{12}.

**Proton pump inhibitor (PPI) trial**

Although pragmatic, symptomatic response to PPI therapy does not equate to a GERD diagnosis, exhibiting an imperfect correspondence with objectively defined disease. On average, 69% of esophagitis patients, 49% of non-erosive reflux disease (NERD) patients, and 35% of patients with normal endoscopy and pH-metry gain symptom relief from a PPI trial\textsuperscript{15}. Hence, when evaluated as a diagnostic test for GERD among heartburn patients, an empiric PPI trial has a sensitivity of 71% and specificity of only 44% compared to the combination of endoscopy and pH-metry\textsuperscript{12,16}. With atypical symptoms (chest pain, chronic cough, laryngitis, etc), PPI response rates are much lower than with heartburn, thereby diminishing the utility of that approach to diagnosis\textsuperscript{17}. A major limitation of the “PPI test” is the strong modulation of symptoms by esophageal hypersensitivity\textsuperscript{18,19}; there is also variation in PPI dosing and duration of the test\textsuperscript{20}. Nonetheless, despite low specificity and high placebo response\textsuperscript{21}, the empiric PPI treatment approach is less costly than diagnostic testing\textsuperscript{22} and is endorsed by societal guidelines\textsuperscript{4,23} undoubtedly leading to the over-diagnosis of GERD and overuse of PPIs.
**Endoscopy and biopsy**

When putative GERD symptoms do not respond to empiric PPI therapy, upper endoscopy (EGD) is advised both to evaluate for GERD complications and to detect potential alternative diagnoses that might redirect therapy. High-grade esophagitis (LA Grades C or D), Barrett’s esophagus, or peptic stricturing are considered confirmatory evidence for GERD. However, erosive esophagitis is found in only 30% of treatment-naive heartburn patients and in <10% when already taking a PPI. Furthermore, most of that is low-grade, and lower grades of esophagitis, particularly LA Grade A, are nonspecific, found in 5-7.5% of asymptomatic controls. When accurately defined, LA Grade B esophagitis provides adequate evidence for initiation of medical management of GERD, but problems with interobserver variability led an expert panel to conclude that additional pH-metry evidence is requisite prior to pursuing anti-reflux surgery (ARS). Barrett’s esophagus is observed in 5-15% of chronic GERD patients, but histological confirmation is documented in only 50% of these. In summary, EGD findings can be clinically important and specific for GERD, but EGD has a low sensitivity in GERD diagnosis.

The Rome IV consensus recommended esophageal biopsies during EGD to rule out eosinophilic esophagitis. Biopsies may also have value in differentiating NERD (with positive pH-metry) from reflux hypersensitivity, functional heartburn and controls when scored using a structured histopathological protocol evaluating papillary elongation, basal cell hyperplasia, dilated intercellular spaces, intraepithelial inflammatory cells, necrosis and erosions; changes that resolve following adequate GERD therapy. However, histopathological findings can overlap between the groups studied, and are not conclusive of GERD. Identification of dilated intercellular spaces on electron microscopy suggests mucosal injury from reflux, but clinical application is limited. The widespread adoption of histopathological examination for GERD injury is hindered by the cumbersome protocol and need for a dedicated esophageal pathologist.
Ambulatory reflux monitoring

Ambulatory reflux monitoring can provide confirmatory evidence of GERD, in patients with normal endoscopy, atypical symptoms and/or when contemplating ARS. Reflux monitoring demonstrates the consequence of GERD pathophysiology, evident as either excessive esophageal acid exposure time (AET) or reflux episodes, rather than the mechanism by which that occurs. Reflux-symptom association utilizes simple ratios and statistical tests to determine whether reflux episodes co-occur with symptoms, and adds value to ambulatory reflux monitoring. Hence, reflux monitoring can confirm or exclude pathologic GERD, albeit not always conclusively.

The primary outcome of a 24-hour pH-metry study is the AET. Extending recording time to 48 or 96 hours with the wireless pH monitoring system increases the diagnostic yield and test reproducibility, and is particularly useful when a transnasal catheter was not tolerated or yielded a negative result despite high suspicion of GERD. However, wireless pH monitoring is expensive, limiting its availability. Another variation on reflux monitoring is pH-impedance monitoring, which characterizes reflux events with both a pH electrode and a series of impedance electrodes. Since pH-impedance detects all reflux (liquid, gas or mixed) regardless of acidity, and defines the direction of flow, it is considered the gold standard. However, the added yield is limited, the test is not widely available, and the interpretation is laborious.

Reflux monitoring can be done ‘on’ or ‘off’ PPI therapy in patients with persistent and/or atypical symptoms despite PPI therapy. The Lyon Consensus proposes that testing always be performed off therapy to demonstrate baseline AET in ‘unproven GERD’, meaning no (or low grade) esophagitis at endoscopy, and no prior positive pH testing. Testing off therapy is also recommended when done to evaluate for ARS. In contrast, the Lyon Consensus proposes that patients with ‘proven GERD’ (prior LA Grade C or D esophagitis, long segment Barrett’s esophagus, or prior abnormal pH-metry) be evaluated on double-dose PPI therapy to establish correlation between refractory symptoms and reflux. 
episodes and/or to exclude inadequate acid suppression or poor compliance as the mechanism of persisting symptoms. This assessment requires pH-impedance (versus pH) monitoring since most reflux episodes on PPI therapy are weakly acidic (pH 4-7)\textsuperscript{49, 50}.

Assessment of proximal esophageal or pharyngeal reflux has also been proposed, but methodology and interpretation have not been standardized, and outcome studies are currently lacking\textsuperscript{51-54}. An additional problem with the pharyngeal pH probe designed to evaluate both aerosolized and liquid acid reflux\textsuperscript{55} is that it detects pharyngeal pH drops in the absence of concomitant esophageal pH-impedance events\textsuperscript{56, 57}, even in post-gastrectomy patients, raising questions about its accuracy\textsuperscript{58}.

\textit{Interpretation of pH and pH-Impedance Monitoring}

Among the pH monitoring metrics, AET is the most reproducible\textsuperscript{59}, is reliably extracted from automated analysis, and is predictive of response from medical and surgical reflux therapy\textsuperscript{60, 61}. However, the significance of an abnormal AET is proportionate to the degree of abnormality, and the Lyon Consensus proposes that AET less than 4% be considered definitively normal (physiologic) and greater than 6% be considered definitively abnormal\textsuperscript{6} with intermediate values between these limits being inconclusive. Another outcome metric of pH-impedance monitoring is the number of reflux episodes (acidic, weakly acidic or weakly alkaline) with the caveat that this is overestimated by the automated analysis and requires manual review of the tracing\textsuperscript{62}. The Lyon GERD Consensus proposes that > 80 reflux episodes per 24 hours are definitively abnormal, while a number <40 is physiologic\textsuperscript{6} and intermediate values inconclusive. The clinical relevance of an abnormal number of reflux episodes remains incompletely defined, although recent preliminary data demonstrate improvement of regurgitation verified by increased reflux episodes following magnetic sphincter augmentation\textsuperscript{63}. Consequently, this is considered an adjunctive measure to be utilized when AET is inconclusive (i.e. between 4-6%). pH-impedance monitoring assists diagnosis of belching disorders and rumination, which
can mimic reflux disease. Additional impedance parameters such as bolus exposure, baseline impedance and post reflux swallow-induced peristaltic wave (PSPW) also have potential as reflux metrics, but outcome data are currently limited.

*Reflux-Symptom Association*

Both pH monitoring and combined pH-impedance monitoring provide analysis of the temporal association between symptoms with a crisp onset (e.g. heartburn, regurgitation, chest pain, cough or belching) and reflux episodes. The time window applied for reflux-symptom association analysis is 2 minutes. The Symptom Index (SI) is the percentage of symptom events preceded by reflux episodes, and the optimal SI threshold for heartburn is 50%. The disadvantage of the SI is that the number of reflux episodes is not considered leaving open the possibility of chance association. The Symptom Association Probability (SAP) and the Ghillebert Probability Estimate (GPE), also known as Binomial Symptom Index (BSI), use more complex statistical calculations to express the probability that symptom events and reflux episodes are associated and are considered positive if the probability (p value) of the observed association occurring by chance is < 5%. Both SAP and GPE/BSI take all relevant components, i.e. total numbers of symptom events, reflux episodes and reflux-related symptom events, into account. In summary, SI is a measure of “effect size”, whereas SAP is a measure of probability. As such the two metrics are complementary, measure different things and cannot be compared to each other. The combination of a positive SI and positive SAP provides the best evidence of a clinically relevant association between reflux episodes and symptoms. Both the SI and SAP are predictive of the effect of medical and surgical anti-reflux therapy, independent of AET.

The reliability of reflux-symptom association analysis is critically dependent on proper execution of the reflux monitoring procedure and meticulous analysis protocols including careful selection of symptoms of interest. Patients must be instructed to use the symptom event button on the portable data logger accurately and to fill in the symptom diary accurately. The outcome of symptom association
analysis is more reliable when at least three symptom events occur during the test. When cough is the symptom of interest, an automated acoustic or manometric cough monitor is necessary to accurately capture cough events. Prolonged wireless pH monitoring increases the yield of symptom association analysis, as does combined pH-impedance monitoring with the detection of weakly acidic reflux episodes. This requires manual analysis, as automated analysis significantly over-detects weakly acidic reflux episodes and inaccurately reports association with non-acid reflux events in nearly 20% of cases. However, “rapid” visual analysis limited to the 2-minute window preceding each symptom event yields SI and SAP values concordant with a full visual analysis, with excellent intra- and inter-observer agreement. Although pH-impedance monitoring provides analysis of symptom-reflux association on PPI therapy, testing off PPI therapy increases the number of symptoms reported, which increases the chance of a positive symptom-reflux association.

Reflex-symptom association analysis has a high degree of reproducibility, the SI being somewhat less reproducible than the SAP. Monte Carlo simulations have identified limitations of reflux-symptom association analysis, especially when acid exposure is low and symptom events are few. If new methods for reflux-symptom association are developed in the future, rigorous outcome testing will be needed to define superiority to the existing metrics.

**Novel Metrics**

Two novel impedance-detected parameters, the PSPW index and baseline impedance have been investigated within GERD phenotypes. These metrics may augment the diagnostic value of impedance-pH testing, especially in discriminating patients with GERD from those with functional heartburn. In health, reflux episodes trigger primary peristalsis to neutralize acidified esophageal mucosa with saliva. This is evident as the antegrade progression of impedance decline within 30 seconds of a reflux episode (PSPW) on a pH-impedance study. The PSPW index, which currently requires cumbersome manual calculation as it is not programmed into the analysis software, consists of the
proportion of reflux episodes on pH-impedance monitoring followed by a PSPW. The PSPW index reflects the integrity of primary peristalsis stimulated by reflux episodes, correlates with contraction reserve assessed using multiple rapid swallows (MRS)\(^9\), and has excellent performance characteristics in differentiating erosive esophagitis and pathological acid exposure from functional heartburn and controls (sensitivity 99-100%, specificity 92%)\(^8\),\(^9\).

Baseline impedance values reflect the permeability of the esophageal mucosa, both in animal models and healthy volunteers, with lower values found in erosive and non-erosive GERD\(^9\),\(^2\),\(^3\). Low baseline esophageal mucosal impedance has been linked to alteration in intercellular space and tight junctions\(^9\),\(^4\),\(^5\) and to reflux symptoms\(^9\). Since frequent swallows and reflux events impact measurement, baseline impedance is best measured from pH-impedance tracings during sleep, termed mean nocturnal baseline impedance (MNBI) when averaged from three 10-minute periods spaced an hour apart\(^8\). MNBI is lower in persisting erosive esophagitis compared to healed esophagitis\(^9\), in PPI-responsive NERD and chronic cough compared to PPI-refractory states\(^8\),\(^5\),\(^8\),\(^9\), and in erosive esophagitis, NERD and reflux hypersensitivity compared to functional heartburn and healthy controls\(^8\),\(^7\),\(^8\),\(^4\). Low MNBI (<2292 ohms) independently predicts response to antireflux therapy\(^9\), links PPI responsive heartburn to reflux better than AET\(^9\), and improves with healing esophagitis\(^1\),\(^0\),\(^0\).

Baseline impedance can also be directly measured from the esophageal mucosa. Initially, this was done using probes with impedance sensors that were passed through the endoscope. Subsequently, the design has improved using two radial sensors mounted on a 10-cm balloon that is inflated to insure optimal contact with a long segment of esophageal mucosa. Baseline mucosal impedance values correlate with esophageal mucosal inflammation, differentiating erosive and non-erosive GERD from eosinophilic esophagitis and normal patients with better specificity (95% vs. 64%) and positive predictive value (96% vs. 40%) compared to pH monitoring\(^1\),\(^2\),\(^3\),\(^4\). Similar to baseline impedance from pH-impedance monitoring, values normalize following PPI therapy. Although normative values are not yet available, ongoing studies should clarify the role of mucosal impedance measurements in GERD management.
**Esophageal high-resolution manometry**

A common indication for high-resolution manometry (HRM) is to accurately place pH or pH-impedance catheters. HRM is also used to assess peristalsis and to detect alternative major motor disorders prior to ARS or when symptoms do not improve with GERD therapy. Consequently, HRM studies are often performed in the setting of GERD. Nonetheless, although fundamental to GERD pathophysiology, a pathophysiologic classification of motor findings in GERD was only recently described7.

**EGJ Barrier Function**

The most fundamental abnormality in GERD is incompetence of the EGJ as an antireflux barrier, making quantifying EGJ competence an attractive biomarker. However, the EGJ is a complex sphincter composed of both the crural diaphragm (CD) and lower esophageal sphincter (LES), the relative dominance of which varies with circumstance. Not only does EGJ pressure vary with time, respiration and swallowing, but EGJ morphology can also vary over time, transitioning between superimposed and separated CD and LES elements105. Furthermore, some degree of EGJ incompetence is physiological, evident by the phenomenon of transient LES relaxation (TLESR), reflex relaxation of both the LES and CD that facilitates gas venting from the stomach106. Clearly, there are challenges to quantifying EGJ barrier function.

No single HRM metric adequately summarizes EGJ competence. Hence, the Lyon Consensus proposes adopting two metrics, one expressing the anatomical morphology of the EGJ and the second summarizing its contractile vigor. EGJ morphology, defined by the relationship between the LES and CD, has been characterized into three subtypes on HRM107, 108: type 1 with superimposed LES and CD, type 2 with axially separated LES and CD pressure signals separated by less than 3 cm, and type 3 with a ≥3 cm
separation between the LES and CD pressure signatures (Figure 1). Type 3 EGJ morphology is associated with reduced LES pressure and lower inspiratory augmentation, which correlates with reflux severity\(^{107, 109, 110}\). However, this relationship is not linear and there are clear exceptions.

The second HRM metric to quantify EGJ barrier function is the EGJ contractile integral (EGJ-CI) (Figure 2). The EGJ-CI is calculated using methodology analogous to that for calculating the distal contractile integral (DCI) with the DCI box set to encompass the LES and CD over a period of three respiratory cycles above a threshold of gastric pressure. The calculated “DCI” is then divided by the duration of the three respiratory cycles to make it independent of time and expressed in units of mmHg•cm\(^{111}\). Several groups of investigators have subsequently tested the performance of the EGJ-CI in segregating GERD populations (Table 1)\(^{111-115}\), with general agreement that this metric identifies a subset of patients with severe barrier dysfunction prone to either endoscopic esophagitis or unequivocally abnormal reflux testing. However, it is also evident from the spread of normal ranges reported among these studies that there are likely methodological discrepancies in exactly how the EGJ-CI is calculated. In view of this, the Lyon Consensus concluded that the EGJ-CI is a promising metric, but needs further research before widespread adoption. In an attempt to standardize methodology among groups, they recommended exclusion of CD component of the EGJ in instances of type 3 EGJ morphology and calculation of EGJ-CI above the gastric baseline pressure.

Further insight into the genesis of the EGJ-CI has been gleaned through studies using 3D-HRM\(^{116}\). Isolation of the CD component of the composite EGJ signal on 3D-HRM concluded that approximately 85% of overall EGJ contractility was attributable to the CD\(^{117}\). Further, analysis of 3D-HRM recordings differentiating the CD and LES constituents of the EGJ pressure complex demonstrated that the CD component correlated strongly with the EGJ-CI, suggesting that both are largely determined by CD contractility\(^{118}\). Together, these studies provide physiological support for adopting the EGJ-CI as a good summary metric of EGJ barrier function, albeit with the caveat that the metric is largely an indicator of CD contractility.
**Esophageal Peristaltic Function**

Esophageal peristalsis can be characterized by the DCI which summarizes the vigor of post transition zone contraction (Figure 2)\(^{108}\). A DCI threshold of 450 mmHg cm s correlates with an averaged distal peristaltic amplitude of 30 mmHg\(^{119}\), the original manometric threshold defining ineffective swallows. When abnormal, esophageal peristalsis is often weak in GERD\(^{120-122}\), with poor and/or delayed formation of post transition zone contraction segment\(^{123}\). This can result in major breaks (>5 cm) in the peristaltic contour even when contraction vigor is preserved, a condition termed fragmented peristalsis when ≥50% of test swallows demonstrate this finding\(^{108}\).

Peristaltic dysfunction becomes progressively more common going from NERD to erosive esophagitis, to Barrett’s esophagus\(^{124, 125}\). High proportions of ineffective contractions increase the likelihood of abnormal AET, particularly while supine\(^{126}\), and increase the likelihood of reflux symptoms\(^{127}\). The Chicago Classification defines ineffective esophageal motility (IEM) as ≥50% of test swallows with DCI<450 mmHg cm s, inclusive of any combination of weak (DCI 100-450 mmHg cm vs) or failed (DCI<100 mmHg cm s) sequences\(^{108}\). Failed sequences are more predictive of an abnormal AET than a similar proportion of weak sequences\(^{128}\). The greatest reflux burden is seen with absent contractility (100% of test swallows with DCI<100 mmHg cm s)\(^{129}\).

**Provocative Tests**

The physiological phenomenon of deglutitive inhibition is more pronounced with multiple swallows in rapid succession such that the esophagus remains in inhibition until after the final swallow, which is then followed by a peristaltic contraction\(^{130}\). Multiple rapid swallows (MRS) and rapid drink challenge (RDC) are two provocative tests of the integrity of deglutitive inhibition during HRM\(^{131}\). With
MRS, five 2-ml swallows are taken less than 4 seconds apart and with RDC 200 ml of water is swallowed within 30 seconds. The Lyon Consensus proposes that every HRM study should be accompanied by at least one of these provocative tests.

Post-MRS contractions are an indicator of ‘contraction reserve’ in the esophagus, the phenomenon wherein the post-MRS contraction has greater DCI than the preceding test swallows (Figure 2). Recent data suggests three MRS sequences for reliable assessment of contraction reserve. The absence of contraction reserve in IEM is predictive of the poor efficacy of promotility drugs, higher AET in NERD, outlet obstruction and subsequent benefit from dilation following ARS, and persistence or development of IEM after ARS. Absent contraction reserve is also the most common manometric finding in systemic sclerosis. The Lyon Consensus accepted the value of adopting MRS into HRM protocols for determining contraction reserve in IEM or absent contractility, acknowledging that MRS is the most widely studied provocative test, provides a computationally simple endpoint (peristaltic augmentation ratio: post-MRS versus pre-MRS), and is quick and easy to perform.

In contrast to MRS, the most important clinical application of RDC is in distinguishing EGJ obstruction from achalasia, by identifying LES relaxation in the former, and an exaggerated pressure gradient across a non-relaxed EGJ in the latter. Therefore, RDC is most helpful in detecting pansophageal pressurization in achalasia, identifying increased resistance to EGJ outflow, and uncovering latent hypercontractility. Additionally, RDC may offer supportive evidence for erosive GERD; effective post-RDC peristalsis was seen in 83% of healthy controls compared to 70% of patients with NERD and only 30% of patients with erosive esophagitis. Solid test meals have also been utilized as provocative tests during HRM, mainly in evaluating transit symptoms. Normal values of esophageal pressure responses to RDC and solid meals have been recently reported in normal healthy volunteers.
HRM studies performed during the post-prandial period could be of interest for identification of pathophysiological mechanisms in GERD\textsuperscript{148, 149}, particularly reflux episodes that tend to be post-prandial. A reflux episode may occur during a TLESR, from low LES pressure, or in conjunction with rumination (increased gastric pressure with or without decreased thoracic pressure) or supragastric belching (air swallowing to initiate belching)\textsuperscript{148}. Post-prandial HRM has also been used to evaluate the efficacy of drugs targeted to TLESRs, rumination, and supragastric belching\textsuperscript{150, 151}. However, several limitations exist, including lack of normative postprandial HRM data, difficulties with standardization of the test meal, and unclear optimal duration of the recording period.

Classification of Motility Findings in GERD

The most common motility pattern in GERD is a normal study. However, either the EGJ or the esophageal body, or both can be abnormal. The EGJ can be hypotensive, with or without a hiatus hernia. Peristalsis can be fragmented, ineffective or absent, with or without contraction reserve. The Lyon Consensus endorses the hierarchical classification of motility findings in GERD first evaluating EGJ morphology and function with LES-CD separation and the EGJ-CI, secondly characterizing the integrity of peristalsis as normal, weak, fragmented or absent, and third, evaluating for contraction reserve\textsuperscript{7} (Table 2). This classification is intended to be used in conjunction with the Chicago Classification.

ADVANCES SINCE THE PORTO CONSENSUS

The Lyon Consensus builds on the Porto Consensus of 2002\textsuperscript{46}, providing recommendations for the use and interpretation of reflux testing techniques in 2017 including esophageal HRM and baseline impedance measurement that were not widely available in 2002 (Table 3). The primary indication for reflux testing is in distinguishing among patients with pathologic reflux burden, reflux mediated hypersensitivity, and functional syndromes (Table 4)\textsuperscript{152}. The Lyon Consensus attempts to augment this
approach by stratifying the significance of findings into those that are conclusive of pathologic GERD, as opposed to suggestive of the diagnosis (Figure 3). The Lyon Consensus also proposes the concept of ‘borderline’ or inconclusive evidence when additional evidence can sway the final judgment toward or away from GERD. This is an area where novel metrics and diagnostic techniques may prove helpful. Conditions that can mimic GERD, such as achalasia, supragastric belching and rumination syndrome need to be excluded with appropriate testing.

*Optimization of GERD Testing*

GERD symptoms are diverse, response to treatment is variable, pathogenesis is heterogeneous, and mechanistic phenotypes are heavily influenced by hypersensitivity and hypervigilance. Because simple algorithms starting with a PPI trial do not consider these complex phenotypes of GERD, they often lead to inappropriate PPI utilization, delayed diagnosis, and inaccurate diagnoses. The Lyon Consensus opines that the optimal initial testing for PPI non-responders with no prior endoscopic or pH-metry demonstration of GERD is pH- or pH-impedance monitoring done withholding antisecretory therapy. A key potential outcome of that testing is to rule out GERD and to redirect management toward weaning off PPIs, using neuromodulators and/or cognitive behavioral therapy as appropriate. In contrast, optimal testing in poorly responsive patients with a prior demonstration of GERD is the combination of EGD, HRM, and pH-impedance monitoring done on twice-daily PPI therapy. This combination of tests serves both to redirect therapy toward alternative diagnoses and to mechanistically subtype patients in terms of poor clearance, excessive reflux episodes and hypersensitivity (Table 4), each of which can trigger specific management options. The precise roles of baseline impedance, PSPW index and provocative maneuvers on HRM remain to be clarified with future research.
**Outcome Measures in GERD**

The optimal use of diagnostic testing may translate into better therapeutic outcomes, but appropriate outcome measures are necessary to properly evaluate that improvement. Esophagitis healing is a common measure for therapeutic trials, but visible esophagitis is rare in patients with refractory GERD symptoms\(^25\), and the objective of the evaluation is to determine if refractory symptoms are attributable to GERD or not. Hence, potentially relevant outcome measures are symptom description\(^154\), symptom questionnaires\(^155\), pictograms with visual depictions of symptoms\(^156,157\), and ambulatory reflux monitoring with analysis of reflux-symptom association\(^75,158\). Outcome measures typically utilized in GERD therapeutic trials have included individual symptom assessment with Likert or visual analog scales\(^159\), global outcome evaluations on Likert scales\(^160\), adequate vs. inadequate relief, disease-specific questionnaires\(^161\), and quality of life questionnaires. Looking to the future, regulatory agencies have stipulated that validated patient reported outcome questionnaires (PROs) and quality of life questionnaires will be requisite in future therapeutic trials. Such PROs require a fastidious approach that includes item generation, testing for reliability, responsiveness, validity and interpretability, and finally, cross-cultural adaptation when applicable, for each diagnostic category\(^162\).

**CONCLUSIONS AND FUTURE DIRECTION**

GERD is a complex disease with a heterogeneous symptom profile and a multifaceted pathogenic basis that defies a simple diagnostic algorithm or categorical classification. The Lyon Consensus defines parameters on esophageal testing that conclusively establish the presence of GERD and characteristics that rule out GERD. Additional evidence from reflux-symptom association, motor findings on HRM, novel metrics from pH-impedance monitoring, baseline mucosal impedance and PPI response complement esophageal testing when pH-metry is borderline or inconclusive. While acknowledging the
limitations of currently available esophageal testing in GERD, the Lyon Consensus proposes this model as a guide to direct management.

The future approach to phenotyping GERD patients should focus on assessing important physiologic biomarkers and PROs to categorize patients based on the severity of refluxate exposure, mechanism of reflux, effectors of clearance, and underlying EGJ pathophysiology (Table 4), while recognizing that no single approach is perfect. Novel metrics assessing tissue resistance, esophageal clearance, peripheral and central neural integration, and psychometrics will allow for a tailored therapeutic approach including pharmacologic treatments, surgical/endoscopic interventions, and behavioral strategies targeting the underlying defect(s) in the antireflux barrier, esophageal clearance, visceral sensitivity, and cognitive response to reflux. As newer metrics emerge, the Lyon Consensus plans future meetings to update and adapt the Consensus conclusions. Collaboration between high volume medical centers involved in GERD testing has opened possibilities for more robust normative data and for validation of conclusions and recommendations from the Lyon Consensus. As the GERD diagnostic paradigm evolves, utilizing diagnostic testing to define a precision approach tailored to the individual patient becomes possible. The goals of evaluation should therefore transition toward defining GERD phenotypes to facilitate tailored treatment.
ACKNOWLEDGEMENT

The authors wish to acknowledge the International Working Group for Gastrointestinal Motility and Function for initiating the consensus meetings and for providing material support for the consensus process.
REFERENCES


Table 1 Studies that have compared the EGJ-CI among patient and control populations. Values reported at median (IQR).

<table>
<thead>
<tr>
<th>Study</th>
<th>Subject groups</th>
<th>EGJ-CI (mmHg•cm)</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nicodème 2014</td>
<td>Controls (n=75)</td>
<td>39 [25-55]</td>
<td>GERD had +++ abnormal pH-impedance studies vs partial + or - for functional</td>
</tr>
<tr>
<td></td>
<td>GERD (n=7)</td>
<td>18* [8-30]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Functional (n=45)</td>
<td>27 [17-69]</td>
<td></td>
</tr>
<tr>
<td>Tolone 2015</td>
<td>Functional (n=39)</td>
<td>22 [10-41]</td>
<td>GERD or functional by endoscopy and pH-impedance testing</td>
</tr>
<tr>
<td></td>
<td>GERD (n=91)</td>
<td>11* [3-21]</td>
<td></td>
</tr>
<tr>
<td>Jasper 2016</td>
<td>Controls (n=65)</td>
<td>63 [50-90]</td>
<td>GERD by pH-metry</td>
</tr>
<tr>
<td></td>
<td>GERD (n=116)</td>
<td>50* [28-70]</td>
<td></td>
</tr>
<tr>
<td>Wang 2016</td>
<td>Controls (n=21)</td>
<td>35 [26-58]</td>
<td>GERD patients underwent fundoplication</td>
</tr>
<tr>
<td></td>
<td>GERD (n=68)</td>
<td>30* [15-53]</td>
<td></td>
</tr>
<tr>
<td>Xie 2017</td>
<td>Controls (n=21)</td>
<td>63 [38-83]</td>
<td>Patients differentiated by pH-impedance and symptom correlation</td>
</tr>
<tr>
<td></td>
<td>Esophagitis (n=39)</td>
<td>22* [20-31]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NERD (n=38)</td>
<td>26* [15-38]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Hypersensitive (n=21)</td>
<td>30* [19-44]</td>
<td></td>
</tr>
<tr>
<td>Ham 2017</td>
<td>Controls (n=23)</td>
<td>67 [27-79]</td>
<td>Non-GERD patients had negative pH-impedance studies</td>
</tr>
<tr>
<td></td>
<td>Esophagitis (n=25)</td>
<td>28* [4-63]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NERD (n=16)</td>
<td>26* [15-32]</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Non-GERD ((n=91)</td>
<td>51 [3-153]</td>
<td></td>
</tr>
</tbody>
</table>

*p<0.05 vs controls or comparator

Methods of EGJ-CI computation were not uniform between these studies, and this might explain differences in calculated thresholds. NERD: non-erosive reflux disease
Table 2. Classification of motor function in GERD using esophageal high-resolution manometry

<table>
<thead>
<tr>
<th>Metrics</th>
<th>Description</th>
</tr>
</thead>
</table>
| **EGJ barrier function**       | **Morphology**                                                                                                   Type 1: superimposed LES and CD  
Type 2: axially separated LES and CD pressure signals separated by less than 3 cm  
Type 3a: ≥ 3 cm separation between the LES and CD pressure signatures with respiratory inversion point at the level of the CD  
Type 3b: ≥ 3 cm separation between the LES and CD pressure signatures with respiratory inversion point at the level of the LES  
**Vigor**                                                                                                   EGJ-CI (mmHg cm)  
DCI box set to encompass the LES and CD over a period of three complete respiratory cycles above a threshold pressure of the gastric baseline  
**Esophageal body motor function**                                                                                                         Distal contractile integral, DCI (mmHg cm s)  
Intact: ≥ 50% of contractions with DCI > 450 mmHg cm s and no defect  
Fragmented: ≥ 50% of contractions with DCI > 450 mmHg cm s and defect > 5 cm  
Ineffective esophageal motility: ≥ 50% of contractions with DCI < 450 mmHg cm s  
Absent peristalsis: 100% of contractions with DCI < 100 mmHg cm s  
**Provocative tests**                                                                                                      MRS (5 liquid swallows - 2ml each - taken less than 4 seconds apart)  
Contractile response  
Failure of contractile response  
Post MRS DCI augmentation  
Absent post MRS contraction  
RDC (free water drinking of 200ml of water within 30 seconds)  
Panesophageal pressurization  
LES relaxation  
Effective post RDC contraction  

Table 3: Comparison of the Porto and the Lyon consensus conclusions.

<table>
<thead>
<tr>
<th>Porto Consensus</th>
<th>Lyon consensus</th>
</tr>
</thead>
</table>
| **No discussion of endoscopy**                                                  | Conclusive endoscopic criteria for GERD  
- LA Grade C or D esophagitis  
- Biopsy-proven Barrett’s esophagus  
- Peptic stricture  
Esophageal impedance monitoring is the only recording method that can achieve high sensitivity for detection of all types of reflux episodes while pH monitoring is required for characterization of reflux acidity. However, the role of impedance monitoring in the management of patients with GERD still needs to be defined.  
**pH-impedance monitoring is the gold standard for detection and characterization of reflux episodes but is expensive, not widely available and interpretation is time consuming.**  
When reflux monitoring is indicated on PPI, pH-impedance should be performed.  
When reflux monitoring is indicated off PPI, the choice between catheter based pH-monitoring, wireless pH monitoring and pH-impedance monitoring is dependent of cost and availability.  
Reflux monitoring is recommended off PPI in instances of “unproven” GERD and on PPI in instances of “proven GERD” (previous LA grade C or D esophagitis, biopsy proven Barrett’s esophagus, peptic stricture or AET off PPI > 6%)  
**An AET < 4% is normal and an AET > 6% is abnormal (whatever the type of reflux monitoring and whether the study was performed off or on PPI)**  
Reflux episodes > 80 / 24 hours is abnormal and < 40 is physiologic on pH-impedance performed off or on PPI. Number of reflux episodes is an adjunctive metric to be utilized when AET is borderline or inconclusive  
**Measurement of baseline mucosal impedance (using either through the scope device or MNBI during ambulatory pH-impedance monitoring) is an adjunctive metric for the diagnosis of GERD**  
A combination of a positive SI and positive SAP provides the best evidence of clinically relevant association between reflux episodes and symptoms. |
| **No discussion of the conditions (off or on PPI) to perform reflux testing**    |                                                                                                                                                                                                           |
| **No discussion of normal values**                                              |                                                                                                                                                                                                           |
| **No discussion of normal values**                                              |                                                                                                                                                                                                           |
Using manometry, common cavities occur during a higher proportion of reflux episodes in neonates and infants than in adults

No discussion of esophageal motor function in GERD

Using manometry, common cavities occur during a higher proportion of reflux episodes in neonates and infants than in adults

No discussion of esophageal motor function in GERD

Esophageal high-resolution manometry is not useful for the direct diagnosis of GERD but can provide adjunctive information:
- to assess EGJ barrier function including its morphology (type I to III) and its vigor (using EGJ-CI)
- to evaluate esophageal body motor function (intact, ineffective, fragmented or absent contractility) that correlates with esophageal reflux burden

Adjunctive tests should be included in the HRM protocol:
- to evaluate the contractile response (multiple rapid swallow)
- to evaluate EGJ obstruction (rapid drink challenge test)

Bilitec™ is a monitoring system that can detect duodeno-gastro-oesophageal reflux (DGOR) by utilizing the optical properties of bilirubin

Bilitec™ is a monitoring system that can detect duodeno-gastro-oesophageal reflux (DGOR) by utilizing the optical properties of bilirubin

Bilitec™ is no longer considered a reliable diagnostic tool for GERD and was not discussed

Bilitec™ is no longer considered a reliable diagnostic tool for GERD and was not discussed
Table 4. GERD phenotypes predicting abnormal reflux burden from clinical evaluation and esophageal testing

<table>
<thead>
<tr>
<th></th>
<th>Pathologic GERD</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High likelihood</td>
</tr>
<tr>
<td><strong>CLINICAL PHENOTYPES</strong></td>
<td></td>
</tr>
<tr>
<td>Symptoms</td>
<td>Heartburn, acid regurgitation</td>
</tr>
<tr>
<td>Endoscopy</td>
<td>High-grade esophagitis, Barrett’s mucosa, peptic stricture</td>
</tr>
<tr>
<td>ROME IV</td>
<td>NERD (abnormal pH metry)*</td>
</tr>
<tr>
<td>Lyon Consensus*</td>
<td>Conclusive evidence of GERD</td>
</tr>
<tr>
<td><strong>MECHANISTIC PHENOTYPES</strong></td>
<td></td>
</tr>
<tr>
<td>Pattern of reflux</td>
<td>Increased acid exposure ± increased numbers of reflux episodes*</td>
</tr>
<tr>
<td>Mechanism of reflux</td>
<td>TLESR Hypotensive EGJ Abnormal EGJ morphology</td>
</tr>
<tr>
<td>Clearance of refluxate</td>
<td>Absent contractility Hiatus hernia</td>
</tr>
<tr>
<td>Cognition, perception of sensation</td>
<td>Appropriate symptom perception, symptom reflux association</td>
</tr>
</tbody>
</table>

*as described by the Lyon Consensus, Figure 3

NERD: nonerosive reflux disease; EGD: esophagogastroduodenoscopy; TLESR: transient lower esophageal sphincter relaxation; EGJ: esophagogastric junction; PSPW: post-reflux swallow induced peristaltic wave
FIGURE LEGENDS

Figure 1. Esophagogastric junction morphology as depicted in HRM. With type 1 morphology the crural diaphragm (CD) component, evident during inspiration (I), is completely superimposed of the lower esophageal sphincter (LES) component such that the magnitude of the actual LES pressure is not discernible. With Type 2 morphology, there is partial separation of the LES and CD constituents, but the respiratory inversion point (RIP) remains at the level of the CD, evident by the decrease observed in the LES pressure band during inspiration. Other characteristics of Type 2 morphology are that the LES-CD separation is <3 cm and that the pressure trough between the LES and CD is greater than intragastric pressure. With Type 3 morphology, there is ≥3 cm separation between the LES and CD and the pressure trough between the two is equal to intragastric pressure during expiration (E). However, the RIP remains at the level of the CD in type 3a and elevated to the level of the LES pressure band with 3b. This is evident by the decreases in LES pressure during inspiration in IIIa and increases in LES pressure during inspiration in 3b.

Figure 2. High-resolution manometry metrics used in the motor classification of GERD. The esophagogastric junction contractile integral (EGJ-CI) measures vigor of the EGJ barrier using a software tool that encompasses length and vigor of the EGJ above the gastric baseline. The measurement is made over three respiratory cycles during quiet rest, and corrected for duration of respiration. The distal contractile integral (DCI) measures vigor of smooth muscle contraction taking length, duration and amplitude of contraction into consideration. Following a series of repetitive swallows (multiple rapid swallows, MRS), DCI augments higher than mean DCI from single swallows when there is contraction reserve.

Figure 3. Interpretation of esophageal test results in the context of GERD. Any one conclusive finding provides strong evidence for the presence of GERD. While a normal EGD does not exclude GERD on its own, this provides strong evidence against GERD when combined with acid exposure time (AET) <4%
and <40 reflux episodes on pH-impedance monitoring off PPI therapy. When evidence is inconclusive or borderline, adjunctive or supportive findings can add confidence to the presence or absence of GERD. Histopathology as an adjunctive measure requires a dedicated scoring system (incorporating papillary elongation, basal cell hyperplasia, dilated intercellular spaces (DIS), intraepithelial inflammatory cells, necrosis and erosions) or evidence of DIS on electron microscopy. However, adjunctive findings, particularly histopathology and motor findings in isolation, are not enough to diagnose GERD. HRM: high-resolution manometry; AET: acid exposure time; MNBI: mean nocturnal baseline impedance; PSPWI index: post-reflux swallow-induced peristaltic wave index; EGJ: esophagogastric junction.