pH Monitoring: Is It the Gold Standard for the Detection of Gastroesophageal Reflux Disease?

Thomas R. Hendrix, MD
Department of Medicine, Johns Hopkins University School of Medicine, Baltimore, Maryland, USA

Abstract. Ambulatory, long term (24-h) intraluminal esophageal pH monitoring is the “gold standard” for detection and quantification of gastroesophageal reflux. Is it, however, the “gold standard” for the diagnosis of gastroesophageal reflux disease (GERD)? The answer depends in part on how GERD is defined. Is it to be defined on the basis of symptoms, inflammatory changes in the esophageal mucosa, extent of the exposure of the esophagus to acid or some combination of these factors? Since the correlation between acid exposure of the esophageal mucosa and either symptoms or histologic changes is poor at best, it seems there must be factors in addition to acid exposure that determine the severity of symptoms and histologic damage. One such is the resistance of the individual patient’s mucosa to injury by acid exposure. In view of the above, it is not surprising that no specific value for acid exposure of the esophagus can be equated with the diagnosis of GERD. The addition of the symptom index, the frequency with which symptoms coincide with reflux episodes, has done little to increase the sensitivity and specificity of pH recording in the diagnosis of GERD. Another variable only occasionally considered is the day to day variation in the frequency and duration of acid reflux. Finally, intraesophageal pH recording measures only the intensity of acid exposure but we have no clinical measure of mucosa resistance to acid-induced injury, the other factor contributing to the pathogenesis of GERD.

Key words: pH monitoring—GERD—Heartburn—Acid regurgitation—Esophagitis—Deglutition—Deglutition disorders.

Intraesophageal pH monitoring provides a continuous record of the reflux of acid from the stomach into the esophagus as well as its neutralization and clearing from the esophagus.

Obtaining these records when the patient is ambulatory, eating normal meals, pursuing normal daily activities, and sleeping in his/her own bed provides the only clinically relevant measure of gastroesophageal reflux. The addition of computer technology to ambulatory intraluminal pH recording has made the storage of the large amount of information produced possible and its analysis easy. Hence, long-term (24-h) intraluminal pH monitoring is the “gold standard” for the detection and quantification of gastroesophageal reflux [1–3].

Is it, however, the gold standard for the diagnosis of gastroesophageal reflux disease (GERD)? Before this question can be addressed, some agreement needs to be reached as to what is meant by the diagnosis of GERD. The term is regularly used but infrequently defined. Its definition is fraught with three major difficulties. First, is the disease to be defined in terms of symptoms, the reflux of acid gastric juice into the esophagus, the pathologic changes produced by acid-peptic juice, that is, peptic esophagitis, either singly or in some combination? Although there is no unanimity in the answers to this question, in this discussion GERD is defined as a disorder characterized by symptoms associated with acid gastric reflux and/or reflux-induced pathologic changes in the esophagus. Second, the symptoms, which are generally believed to indicate the presence of GERD, for example, heartburn, acid regurgitation, water brash, noncardiac chest pain, belching, morning hoarseness, nocturnal cough, nonseasonal asthma, are nonspecific, variable, and do not form a predictable constellation of symptoms. Of these symptoms only heartburn and acid regurgitation have been shown to have a useful degree of specificity but they are not very sensitive [4]. To make matters
worse, many patients with the most damaging reflux, for example, those with ulcerative esophagitis, peptic stricture, or Barrett’s esophagus, often have insufficient symptoms to lead them to seek medical attention. In these circumstances it is only when dysphagia due to a benign or malignant stricture appears that reflux disease is recognized [5]. Third, all individuals, whether symptomatic or not, can be shown to have gastroesophageal reflux [3]. In asymptomatic individuals reflux occurs more frequently after meals and uncommonly in the supine position. There is then a problem in defining at what level the frequency and/or duration of reflux episodes are no longer physiologic (normal) and enter the realm of pathologic (disease—GERD).

An intraesophageal pH below 4.0 for more than 4% of the 24-hour recording time is the most commonly used level to divide normal from abnormal reflux [3]. This value was obtained by placing the upper limit of normal 2 standard deviations above the mean of the asymptomatic control group. In most studies, pH is recorded at a point 5 cm above the lower esophageal sphincter, as determined manometrically, or 5 cm above the level where a sudden rise in pH is recorded during slow withdrawal of the pH electrode from the stomach into the esophagus.

Recently, it was pointed out that the distribution of time below pH 4.0 does not follow a normal distribution, hence 2 SDs above the mean is not a statistically defensible value for the boundary between normal and abnormal reflux. A more sophisticated statistical analysis of 45 patients with GERD (clinically defined and 42 controls) indicated that a more appropriate separation is provided by a threshold of 10.5% of the upright portion below pH 4.0 and 6.0% of the supine portion of the 24-h pH recording. Abnormal or pathologic reflux is then defined as being present if either one or both periods (upright and supine) are above threshold. Using these criteria, sensitivity of 93.3% and specificity of 92.9% were reported [7]. Recently, the same research group revised the thresholds downward based on observations made in a larger group of patients (n = 304). The 24-h pH record was classified as abnormal if the pH was below 4.0 for greater than 8.2% of the upright recording or greater than 3.0% of the supine recording time [4].

As ambulatory pH monitoring is usually performed, the temporal relation between “reflux symptoms” to reflux episodes is not quantified. In an attempt to improve the sensitivity and specificity of 24-h intraesophageal pH recording in the diagnosis of GERD, calculation of a symptom index has been suggested. The symptom index was calculated from the frequency with which symptoms, heartburn, chest pain, and the patient’s chief complaint were associated with a recorded reflux episode (within 5 min of a fall in pH to below 4.0). If each reported reflux symptom was associated with a reflux episode as recorded by pH monitoring the symptom index is 100%; if none are, the index is 0%. Values above 75% or below 25% were contributed to accuracy of the diagnosis of GERD (clinically significant gastroesophageal reflux [6]. In a study of the reliability of symptoms in predicting GERD, symptoms in 304 patients, as elicited by an experienced gastroenterologist, were compared with the results of 24-h pH monitoring. The utility of the symptom index in arriving at a diagnosis was also evaluated. The symptom index would add 12% more patients to those diagnosed as GERD by pH monitoring criteria (primarily patients with normal pH 24-h recording but in whom greater than 75% of symptom episodes were associated with recorded reflux episodes). On the other hand, if patients whose dominant symptom was heartburn and/or acid regurgitation were considered to have GERD regardless of pH monitoring, results determining a symptom index added nothing to the “precision” of the diagnosis of GERD [4].

In addition to the absence of a close correlation between reflux symptoms and pH monitor recorded reflux episodes, the day-to-day variation in reflux may also limit the precision of pH monitoring in the diagnosis of GERD. Although it is general experience that the symptoms of reflux vary from day to day, there is little information on the day-to-day or month-to-month variation in 24-h intraesophageal pH recording. In a study of 20 randomly selected patients, intraesophageal, 24-h pH monitoring was conducted on two consecutive days. Reflux was normal or abnormal on both test days in 75% of the subjects. Concordance was greater with upright reflux (83%) than with supine (62%) or total reflux (77%). Discordance did not follow any pattern (e.g., normal on day 1 and abnormal on day 2 [8]). In a differently designed study, 53 subjects (14 asymptomatic individuals, 14 with documented esophagitis, and 25 with atypical or nonspecific reflux symptoms) were studied with 24-h pH recording 10 days apart. Although repeat values for percent recording time below pH 4.0 varied as much as 3.2-fold, there was a 90% chance that the repeat test would be concordant if the initial test was clearly normal (pH <4.0 less than 1.8% recording time) or clearly abnormal (pH <4.0 more than 9.4% of the recording time). The overall reproducibility, normal or abnormal, was 80% using 4.38% of total recording time as the diagnostic endpoint [9]. Although these attempts to develop optimal criteria for separating normal from abnormal reflux and in the process defining the disorder called GERD have clinical utility and statistical validity, they distract from developing an understanding of the pathogenesis of GERD.

Numerous studies have pointed out poor correlation between the frequency and severity of reflux symp-