OPEN ACCESS

GAVIN PUBLISHERS

Journal of Gastrointestinal Disorders

Patel R, et al. J Gastrointest Disord 2: 1008.

DOI: 10.29011/JGID-1008.001008

Research Article

96-Hour Esophageal pH Monitoring: The Tiebreaker for Abnormal DeMeester Score and Symptom Index

Rupal Patel[2], Ambuj Kumar[3], Soojong Chae[2], John Jacobs[1,2], Joel Richter[1,2*]

[1]Joy McCann Culverhouse Center for Esophageal Diseases, University of South Florida College of Medicine, USA

[2]Division of Digestive Diseases and Nutrition, University of South Florida College of Medicine, USA

[3]Program for Comparative Effectiveness Research, University of South Florida College of Medicine, USA

[*]**Corresponding Author:** Joel Richter, Division of Digestive Diseases and Nutrition, University of South Florida, USA. Tel: +1-8139743980; Email: jrichte1@health.usf.edu

Citation: Patel R, Kumar A, Chae S, Jacobs J, Richter J (2019) 96-Hour Esophageal pH Monitoring: The Tiebreaker for Abnormal DeMeester Score and Symptom Index. J Gastrointest Disord 2: 1008. DOI: 10.29011/JGID-1008.001008

Received Date: 10 July, 2019; **Accepted Date:** 14 August, 2019; **Published Date:** 20 August, 2019

Abstract

Background: A commonly utilized diagnostic test for GERD is distal esophageal pH monitoring over 48 hours. It is now possible to record 96 hours (4 days) of pH data with this device using the same lithium battery with a longer life. We addressed three clinical questions: 1) What is the overall additional diagnostic gain from the 48 hour to 96-hour study? 2) What is the diagnostic yield of different days and for different reflux parameters? 3) Is there an improvement in diagnostic yield in the day 1/day 2 concordant and discordant groups?

Methods: All consecutive adult patients (>18 years) undergoing a 96hour wireless capsule study (Bravo; Given Imaging, Atlanta, GA) between March 2016 - October 2017 were eligible for inclusion. Any patient who had pH testing inadvertently done on PPIs, had chest pain requiring removal of the pH capsule prematurely, or had the capsule fall off before 96 hours of data recording was complete were excluded from the study.

Key Results: Day 1 and Day 2 data for DeMeester scores showed 53% of our study population (n=45) was in the +Day 1/+Day 2 DeMeester score group, 27% was in the -Day 1/-Day 2 DeMeester score group, 11% was in the +Day 1/–Day 2 DeMeester score group, and 9% was in the -Day 1/+ Day 2 DeMeester score group. Overall, the 48 hours of additional reflux data changed a diagnosis in 10/45 patients (22%) based on the DeMeester scores. ROC analysis showed that total % time pH<4 and DeMeester score had outstanding accuracy on days 2 through 4. The upright % time pH<4 had outstanding accuracy on days 2 and 3. Day 1 data for all parameters had poor diagnostic accuracy.

Conclusions: Our data shows that the overall diagnostic yield is increased when extending the esophageal pH study to 96 hours, allowing a confident diagnosis in an additional 22% of patients. The first day of data from esophageal pH studies is not consistently accurate, whereas the subsequent days are highly accurate for the most important study variables of DeMeester and % total time pH<4.

Keywords: Ambulatory pH Testing; Bravo pH Testing; DeMeester Score; Gastroesophageal Reflux Disease

Gastroesophageal Reflux Disease (GERD) remains a prevalent and morbid condition in the Western world. However, the diagnosis is not always straightforward and the evaluation for GERD can extend beyond a PPI challenge. A commonly utilized test is distal esophageal pH monitoring over 48 hours [1-3]. This pH monitoring test breaks down pH data into separate 24 hours blocks (day 1 and day 2). For each day, the pH data includes: % total time pH<4, % upright time pH<4, % supine time pH<4, symptom index, symptom association probability, and DeMeester score [4].

Our group recently performed and published a study on the 48 hour distal esophageal pH monitoring system [5]. We found that significantly more acid reflux is captured by the capsule on the first day and when the capsule is placed in the afternoon. The cause was uncertain and our recommendation was to place the capsule in the morning whenever possible. We also observed that about 25% of patients had discordant DeMeester scores between day 1 and day 2 leading to confusion in making the diagnosis of GERD. It is now possible to record 96 hours (4 days) of pH data with this

device using the same lithium battery with a longer life, increasing the amount of acid reflux data available for each patient. We now perform 4 day studies regularly at our Swallowing Center. This technology is friendly and useful. The placement of the capsule in the distal esophagus via EGD is the same, the patient only need return to clinic after 4 days instead of 2 days to download data. We believe the longer study helps to better clarify the severity and frequency of GERD, helps to better correlate with activities such as exercise and may help resolve the frequent dilemma of the discordant 2day study.

Our study is a retrospective cohort study involving patients who have undergone the 96 hour distal esophageal pH study. This has given us the opportunity to address three clinical questions:

- What is the overall additional diagnostic gain from the 48 hour to 96 hours study?
- What is the diagnostic yield of different days and for different reflux parameters?
- Is there an improvement in diagnostic yield in the day 1/day 2 concordant and discordant groups?

Materials and Methods

Study Design

We performed a retrospective cohort study. The institutional review board at the University of South Florida approved this study (PRO00015681). Patient data were deidentified throughout the collection process to ensure patient privacy in accordance with the Health Insurance Portability and Accountability regulations.

Participants

All consecutive adult patients (>18 years) undergoing a 96 hours wireless capsule study (Bravo; Given Imaging, Atlanta, GA) between March 2016 through October 2017 were eligible for inclusion. Any patient who had pH testing inadvertently done on PPIs, had chest pain requiring removal of the pH capsule prematurely, or had the capsule fall off before 96 hours of data recording was complete were excluded from the study.

The 96 Hour Wireless Capsule Study Procedures and Setting

Patients with suspected GERD seen at our Swallowing Center who undergo the 96 hour's wireless capsule study is all prescreened and should meet at least one of the following criteria: have "Intractable" GERD symptoms, complain of atypical GERD symptoms, and/or are presenting for evaluation before antireflux surgery. All studies are performed after an overnight fast with the patient off antireflux therapy (proton pump inhibitors discontinued 1 week prior, histamine-2 receptor antagonists discontinued 3 days prior, antacids discontinued the night prior).

Three gastroenterologists (JR,SC,JJ) familiar with the technique placed all capsules. Before placement, calibration (pH 1 and 7) was done in the Swallowing Center by experienced nurses who assisted with the capsule placement. Upper endoscopy was performed under propofol anesthesia. After complete examination of the upper GI tract, the distance between the squamocolumnar junction and the incisors was measured. The endoscope was removed and the pH monitoring device deployed blindly (6 cm above the squamocolumnar junction) using the Bravo delivery system guided by the measurements obtained from endoscopy.

After wireless capsule placement, all patients remained off PPIs, histamine-2 receptor antagonists, and antacid medications for 4 days. They were encouraged to be active: to eat whatever they desired, exercise, go to work, and resume their normal daily routine. All patients kept a diary recording when napping or sleeping and when meals were eaten. They were also instructed to record what they ate or drank other than water (these periods were excluded in the analysis). Patients were asked to identify a dominant symptom for symptom analysis. Patients typically chose one of the following symptoms as their dominant complaint: heartburn, regurgitation, or chest pain. Patients were instructed to press the symptom indicator button on the pH recorder when experiencing only their one dominant GERD symptom.

Patients returned to the Swallowing Center 96 hours after capsule placement. At this time, data were downloaded and reviewed by a nurse, then analyzed by a computer separately. The final review and diagnosis was made by the senior author (JR).

Data Collection

We collected data on day 1, day 2, day 3, day 4, and the average total reflux across 4 days for each patient. This data included: % total time pH<4, % upright time pH<4, % supine time pH<4, symptom index, and DeMeester score. Demographic data on age and sex were collected as well. All the above data were stored on a deidentified spreadsheet.

Descriptive Analysis

We calculated the proportion of patients in the following 4 groups: +Day 1/+Day 2 DeMeester scores, -Day 1/-Day 2 DeMeester scores, -Day 1/+Day 2 DeMeester scores, and +Day 1/-Day 2 DeMeester scores. The DeMeester score was considered positive if > 14.72. The DeMeester score was chosen for the primary variable as it is the most reproducible of the commonly analyzed pH variables (similar to percent total time pH<4), and has an accepted cutoff point at 14.72 [6]. Overall pathologic GERD was defined as ≥2 days of a positive DeMeester score or positive symptom index. Hasak, et al. used the same criteria in their recent study [7]. We arbitrarily defined improvement in diagnostic yield as: a diagnosis of GERD was made or excluded based on the 96 hours study when the first 48 hours yielded a negative or discordant study.

Statistical Analysis

Receiver Operating Curve (ROC) analysis was performed of total percent time pH <4, DeMeester score, and symptom index to assess the diagnostic accuracy for GERD for each individual day in the Day 1/Day 2 groups and summarized as Area Under the Curve (AUC) along with 95% confidence intervals.

Results

Participants

A total of 45 patients (62% female, average age 51 years) met the inclusion criteria. Table 1 describes the cohort's major symptoms, with each patient identifying one dominant symptom.

Total Sample Size (n)	45
% male	38
% female	62
Average age (years)	51
Dominant Symptom (n)	
Heartburn	21
Globus sensation	8
Cough	4
Chest pain	4
Regurgitation	4
Hoarse voice	3
Other	1

Table 1: Patient Demographics and Reported Symptoms.

Descriptive Analysis

Day 1 and Day 2 data for DeMeester scores showed 53% of our study population was in the +Day 1/+Day 2 DeMeester score group, 27% was in the -Day 1/-Day 2 DeMeester score group, 11% was in the +Day 1/-Day 2 DeMeester score group, and 9% was in the -Day 1/+ Day 2 DeMeester score group. Day 1 and Day 2 data for Symptom Indices showed 16% of our study population was in the +Day 1/+Day 2 Symptom indices score group, 52% was in the -Day 1/-Day 2 Symptom indices score group, 16% was in the +Day 1/-Day 2 Symptom indices score group, and 16% was in the -Day 1/+Day 2 Symptom indices score group.

Table 2 shows additional information gained from day 3 and 4 in making the diagnosis of GERD and symptom association. Overall, the 48 hours of additional reflux data changed the diagnosis (either supported or excluded a diagnosis of GERD) in 10/45 patients (22%) based on the DeMeester scores. 9/10 (90%) of these patients were in the discordant DeMeester score groups. In the +/- discordant group, 2/5 (40%) patients had negative scores both additional days. In contrast, all patients in the -/+ discordant group had at least an additional positive day and three were positive the remaining two days. The rate of maintaining concordance of the DeMeester groups was high, with 20/24 (83%) of the +Day 1/+Day 2 DeMeester score group maintaining concordance and 11/12 (92%) of the -Day 1/-Day 2 DeMeester score group maintaining concordance in days 3 and 4. As detailed in Table 3, the 48 hours of additional reflux data clarified a diagnosis in 9/25 patients (36%) based on symptom indices.

n=45	**Day 1 and 2 DeMeester scores**		**Days 3 and 4 DeMeester scores**
Concordant			
24	+	+	• 20 patients had positive day 3 and day 4 DeMeester scores • 4 patients had only 1 additional positive day
12	-	-	• 11 patients had negative day 3 and day 4 DeMeester scores • 1 patient had positive day 3 and day 4 DeMeester scores
Discordant			
5	+	-	• 3 patients had only 1 additional positive day • 2 patients had negative day 3 and day 4 DeMeester scores
4	-	+	• 3 patients had discordant day 3 and day 4 DeMeester scores • 1 patient had positive day 3 and day 4 DeMeester scores

Table 2: Additional Information Gained from Day 3 and Day 4 DeMeester Scores.

n=25	Day 1 and 2 Symptom Indices		Days 3 and 4 Symptom Index
Concordant			
4	+	+	• 2 patients had positive day 3 and day 4 symptom indices • 2 patients had only 1 additional positive day
13	-	-	• 1 patient had positive day 3 and day 4 symptom indices • 2 patients had only 1 additional positive day • 10 patients had negative day 3 and day 4 symptom indices
Discordant			
4	+	-	• 1 patient had positive day 3 and day 4 symptom indices • 1 patient had only 1 additional positive day • 2 patients had negative day 3 and day 4 symptom indices
4	-	+	• 3 patients had positive day 3 and day 4 symptom indices • 1 patient had only 1 additional positive day

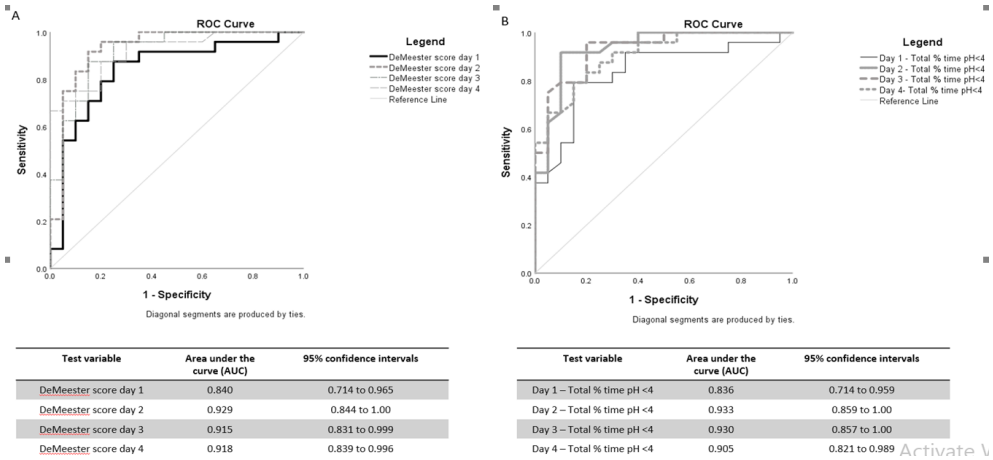
Table 3: Additional Information Gained from Day 3 and Day 4 Symptom Indices.

ROC Analysis

Table 4 summarizes the diagnostic AUC across the 4 days. The total % time pH<4 and DeMeester score paralleled each other and had outstanding accuracy on days 2 through 4 (Figure 1). The upright % time pH<4 had outstanding accuracy on days 2 and 3, but supine % time pH<4 and SI were less accurate. Day 1 data for all parameters had poor diagnostic accuracy.

	Day 1	Day 2	Day 3	Day 4
	AUC (95% confidence intervals)			
Total % time pH<4	0.836 (0.71-0.96)	**0.933 (0.86-1)**	**0.930 (0.86-1)**	**0.905 (0.82-0.99)**
Upright % time pH<4	0.831 (0.71-0.96)	**0.94 (0.86-1)**	**0.90 (0.81-0.99)**	0.87 (0.76-0.97)
Supine % time pH<4	0.74 (0.58-0.89)	0.79 (0.65-0.94)	0.81 (0.67-0.95)	0.83 (0.71-0.96)
DeMeester score	0.84 (0.71-0.97)	**0.929 (0.84-1)**	**0.915 (0.83-1)**	**0.918 (0.84-1)**
Symptom Index	0.643 (0.42-0.86)	0.87 (0.73-1)	0.812 (0.63-0.99)	0.818 (0.63-1)
AUC > 0.90 accuracy are illustrated in bold print.				

Table 4: Area Under the Curve for the Various Acid Reflux Measurements of Days 1 Through 4.



Test variable	Area under the curve (AUC)	95% confidence intervals
DeMeester score day 1	0.840	0.714 to 0.965
DeMeester score day 2	0.929	0.844 to 1.00
DeMeester score day 3	0.915	0.831 to 0.999
DeMeester score day 4	0.918	0.839 to 0.996

Test variable	Area under the curve (AUC)	95% confidence intervals
Day 1 – Total % time pH <4	0.836	0.714 to 0.959
Day 2 – Total % time pH <4	0.933	0.859 to 1.00
Day 3 – Total % time pH <4	0.930	0.857 to 1.00
Day 4 – Total % time pH <4	0.905	0.821 to 0.989

Figure 1: Receiver Operating Characteristics (ROC) analysis for using A) DeMeester Score and B) Total % time <4 to determine acid reflux disease. Day1 data for both these parameters have the worst diagnostic accuracy.

Discussion

As we discovered in our previous study, approximately 25% of patients with similar demographics and symptoms to our study population will have discordant Day 1 and Day 2 DeMeester scores. Our current study implies that in this group with discordant DeMeester scores, the 96-hour study gives us the opportunity to resolve the diagnosis with a greater deal of certainty with an increased diagnostic yield of 22%.

A few recent studies have examined the clinical utility of the 96 hours pH monitoring system. Hasak, et al. [7] evaluated the clinical value of prolonged wireless pH monitoring (96 hours), the data from which is available in abstract form. They describe categorization of acid exposure into dominant patterns based on ≥2 days of pathologic, physiologic, or borderline acid exposure time. Acid exposure time from the first 2 days was concordant in 174/265 patients (65%). This correlated with the dominant pattern from the overall study in 98% of patients ($p<0.01$). When the first 2 days were discordant, the remainder of the study allowed categorization into the dominant pattern in 82% of patients (48 physiologic,31 pathologic,12 borderline, $p<0.05$). If day 1 was borderline, 57% of patients were categorized into the physiologic group and 17% of patients into the pathologic group. This study supports the concept that adding additional days improves diagnosis in borderline or discordant cases.

Scarpulla, et al. [8] performed a retrospective analysis of 83 patients who received a 96 hours' esophageal pH study. Complete four day readings were available for 34/83 (41%) patients. They collected esophageal acid exposure, DeMeester score, and symptom index for their patients. They observed that variation in pH measurements were reduced with increasing study duration and significant improvement in diagnostic sensitivity with study duration ($p<0.01$). When they compared the diagnosis from the three interval test periods to the 96 hour "Gold standard" diagnosis, they found a consistent diagnosis in 22 (63%), 29 (83%), and 32 (91%) patients for the 24 hours, 48 hours, and 72 hours test periods respectively.

Our data shows that the overall diagnostic yield is increased when extending the esophageal pH study by two days to 96 hours, allowing a confident diagnosis in an additional 22% of patients. However, some reflux parameters fared better than others in helping to resolve an uncertain diagnosis. The five parameters uniformly calculated during an esophageal pH study are: total % time pH<4, upright % time pH<4, supine % time pH<4, DeMeester score, and Symptom Index. As illustrated in Table 4, the supine % time pH<4 and symptom index are not worthwhile in a situation where a diagnosis is uncertain. Across all four days, both supine % time pH<4 and symptom index had poor diagnostic accuracy ranging from 0.64-0.87. On the other hand, the upright % time pH<4, the total % time pH<4, and DeMeester scores all give an accurate diagnosis across days 2 and 3 (illustrated by bold numbers in Table 4). The day 4 total % time pH<4 and DeMeester score also yield an accurate diagnosis.

Importantly, this study highlights the variability in the day 1 reflux data obtained by esophageal pH studies. The AUC of day 1 parameters ranged from 0.64 to 0.84 (see Table 4), implying that the diagnostic accuracy of Day 1 reflux testing is not very high. Therefore, if a patient has +Day 1/-Day 2 DeMeester scores, this is not necessarily reliable to make a diagnosis of GERD. As see in Table 3, 2/5 (40%) patients in the +/- discordant group had negative scores both additional days suggesting a high rate of false positive diagnoses with just 2 days of data. The etiology for the great variability in Day 1 data has not been resolved, but likely is due to timing of the Bravo procedure (morning vs afternoon), carry-over anesthesia effect and more sedentary activities post procedure [5].

In conclusion, the first day of data from esophageal pH studies is not consistently accurate, whereas the subsequent days are highly accurate for the most important study variables of DeMeester and % total time pH<4. The 96-hour esophageal pH study is useful because 2 additional days of reflux data improves the confidence in making a GERD diagnosis. With the latest technology, it is now possible to record 96 hours (4 days) of pH data using a single lithium battery. However, even with the older battery, the patient can return to clinic after 2 days where a new battery is placed, easily allowing for collection of 96 hours of data. These additional two days add no risk to the procedures and are greatly appreciated by the patients when a potential diagnostic dilemma is resolved.

References

1. Richter JE, Pandolfino JE, Vela MF, Kahrilas PJ, Lacy BE, et al. (2013) Utilization of wireless pH monitoring technologies: a summary of the proceedings from the esophageal diagnostic working group. Dis Esophagus 26: 755-765.
2. Lacy BE, Dukowicz AC, Robertson DJ, Weiss JE, Teixeira P, et al. (2011) Clinical utility of the wireless pH capsule. J Clin Gastroenterol 45: 429-435.
3. Vaezi MF, Sifrim D (2018) Assessing old and new diagnostic tests for gastroesophageal reflux disease. Gastroenterology 154: 289-301.
4. Hirano I, Richter JE (2007) ACG practice guidelines: esophageal reflux testing. Am J Gastroenterol 102: 668-685.
5. Patel R, Chae S, Kumar A, Richter JE (2017) Sedation and afternoon placement of the 48-hour wireless ambulatory pH capsule results in more reflux on the first day. J Clin Gastroenterol 51: 594-598.
6. Ayazi S, Lipham JC, Portale G, Peyre CG, Streets CG, et al. (2009) Bravo catheter-free pH monitoring: normal values, concordance, optimal diagnostic thresholds, and accuracy. Clinical Gastroenterology and Hepatology 7: 60-67.
7. Hasak S, Yadlapati R, Pandolfino JE, Gyawali CP (2018) Su1076 - clinical value of prolonged wireless pH monitoring in characterizing esophageal acid exposure. Gastroenterology 154: 478.
8. Scarpulla G, Camilleri S, Galante P, Manganaro M, Fox M (2007) The impact of prolonged pH measurements on the diagnosis of gastroesophageal reflux disease: 4-day wireless pH studies. Am J Gastroenterol 102: 2642-2647.