Inconsistency in the Diagnosis of Functional Heartburn: Usefulness of Prolonged Wireless pH Monitoring in Patients With Proton Pump Inhibitor Refractory Gastroesophageal Reflux Disease

Roberto Penagini,1* Rami Sweis,2 Aurelio Mauro,1 Gerson Domingues,3 Andres Vales4 and Daniel Sifrim5

1Gastroenterology and Endoscopy Unit, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, Department of Pathophysiology and Transplantation, Università degli Studi di Milano, Milan, Italy; 2University College London Hospital, London, United Kingdom; 3State University of Rio de Janeiro, Rio de Janeiro, Brazil; 4Esophageal Lab, Department of Gastroenterology, Guy’s & St Thomas’ NHS Foundation Trust, London, United Kingdom; and 5Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, United Kingdom

Background/Aims
The diagnosis of functional heartburn is important for management, however it stands on fragile pH monitoring variables, ie, acid exposure time varies from day to day and symptoms are often few or absent. Aim of this study was to investigate consistency of the diagnosis of functional heartburn in subsequent days using prolonged wireless pH monitoring and its impact on patients’ outcome.

Methods
Fifty proton pump inhibitor refractory patients (11 male, 48 years [range, 38-57 years]) with a diagnosis of functional heartburn according to Rome III in the first 24 hours of wireless pH monitoring were reviewed. pH variables were analysed in the following 24-hour periods to determine if tracings were indicative of diagnosis of non-erosive reflux disease (either acid exposure time > 5% or normal acid exposure time and symptom index ≥ 50%). Outcome was assessed by review of hospital files and/or telephone interview.

Results
Fifteen out of 50 patients had a pathological acid exposure time after the first day of monitoring (10 in the second day and 5 in subsequent days), which changed their diagnosis from functional heartburn to non-erosive reflux disease. Fifty-four percent of non-erosive reflux disease vs 11% of functional heartburn patients (P < 0.003) increased the dose of proton pump inhibitors or underwent fundoplication after the pH test. Outcome was positive in 77% of non-erosive reflux disease vs 43% of functional heartburn patients (P < 0.05).

Received: June 12, 2014 Revised: December 1, 2014 Accepted: December 7, 2014
©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
*Correspondence: Roberto Penagini, MD Gastroenterology and Endoscopy Unit, Università degli Studi di Milano, Fondazione IRCCS Ca’ Granda Ospedale Maggiore Policlinico, via della Commenda 19, 20122 Milan, Italy Tel: +39-02-55034591, Fax: +39-02-50320403, E-mail: roberto.penagini@unimi.it

Financial support: None.

Conflicts of interest: None.

Author contributions: Roberto Penagini had the concept of the study and wrote the paper; Rami Sweis and Andres Vales acquired data; Aurelio Mauro acquired data, made statistical analysis, and wrote the paper; Gerson Domingues acquired data; and Daniel Sifrim had the concept of the study, wrote the paper, and revised the manuscript. All authors approved the final version of the article, including the authorship list.

ORCID: Roberto Penagini, http://orcid.org/0000-0001-6918-9479.
Conclusions
One-third of patients classified as functional heartburn at 24-hour pH-monitoring can be re-classified as non-erosive reflux disease after a more prolonged pH recording period. This observation has a positive impact on patients’ management. (J Neurogastroenterol Motil 2015;21:265-272)

Key Words
Esophageal pH monitoring; Gastroesophageal reflux; Proton pump inhibitors

Introduction
Heartburn is a common symptom experienced by the general population and it is highly suggestive for gastroesophageal reflux disease (GERD).1-3 Non-erosive reflux disease (NERD) is the most common presentation of this illness, affecting up to 70% of individuals.4-5 Patients with heartburn are a heterogeneous group including various subpopulations with different mechanisms responsible for their symptom, ie, patients with typical GERD symptoms and either increased esophageal acid exposure or acid hypersensitivity, and patients with functional heartburn (FH).6 The distinction between patients with GERD related and those with GERD unrelated heartburn (ie, FH), is clinically highly relevant because the first group benefits from pharmacological or surgical anti-reflux therapy, whereas the second one should not receive prolonged unnecessary proton pump inhibitor (PPI) treatment and, even less, antireflux surgery.

Twenty-four hour esophageal pH-monitoring has long been considered the best examination in order to differentiate the various types of heartburn patients because it can establish the amount of esophageal acid exposure and the correlation between reflux and symptoms.7,8 More recently addition of impedance to measurement of pH has enabled detection of patients with heartburn related to non-acid reflux and their separation from those having FH.7 So far studies on pathophysiological differences between patients with FH and those with NERD, for example intercellular space diameter or histopathological findings, basal impedance, acid sensitivity, somatisation, have relied on 24-hour pH or pH and impedance monitoring.10-14

However 24-hour reflux tests have some intrinsic limitation: previous studies have shown that esophageal acid exposure time (AET) presents a physiological day to day variability due to different dietary habits and physical activity among days;15-17 symptoms during single day monitoring are often absent or few, making either symptoms reflux association indexes unreliable or impossible to be evaluated; finally, patients are often uncomfortable and adopt a more sedentary life style, resulting in reduction of reflux-provoking activities,18,19 which may further impair the diagnostic yield of catheter based studies. These limitations could be particularly relevant to patients with NERD who often have normal acid exposure during 24 pH-monitoring, differently from patients with esophagitis who generally have increased reflux20 and could have biased results of studies aimed at investigating differences between FH and NERD.

The Bravo wireless technique has been developed in order to overcome limitations of the traditional catheter-based pH testing. This well-known system uses a pH capsule fixed to the esophageal mucosa that transmits data to a portable receiver using radiotelemetry.21 The wireless test is better tolerated compared with the catheter based one,22-25 usually lasts 48 hours and can be extended further, increasing its ability to record pathological reflux and/or establish symptoms/reflux association.26-29 As a matter of fact prolonged wireless pH-monitoring has been shown to diagnose GERD in patients with a previously negative 24-hour catheter based monitoring.30

Aim of our study was, therefore, to investigate consistency of the diagnosis of FH after the first day of pH recording using 48-96 hours wireless pH monitoring and impact of prolonged pH monitoring on outcome.

Materials and Methods

Study Patients
Fifty consecutive PPI refractory patients (11 male, 48 years [range, 38-57 years]) with heartburn as dominant complaint and a diagnosis of FH according to Rome III,31 ie, no mucosal breaks at upper gastrointestinal endoscopy, normal AET (< 5% of time at pH < 4),16 and symptom index (SI) < 50% in the first 24 hours of wireless pH monitoring, were retrospectively analysed as part of an international collaborative study (Brazil [n = 12], Italy [n = 15], and UK [n = 23]). As there are different definitions for refractoriness to PPIs we have chosen the most in-
clusive one by considering all patients that were refractory to a stable single or double dose PPI treatment. Patients with no available data about PPI response were not included in the study. Antisecretory drugs (PPI) were stopped at least 8 days before pH monitoring; patients were instructed not to use antacid formulations during the whole recordings. Informed written consent was obtained from each patient.

Forty-eight to 96-hour Wireless pH Monitoring

After external calibration in buffer solutions pH 1.0 and 7.0, the Bravo pH capsule (Medtronic Inc, Shoreview, MN, USA) was attached 6 cm above the squamocolumnar junction during upper endoscopy under mild conscious sedation (1-5 mg midazolam i.v.) as previously described.\textsuperscript{23} pH monitoring was then activated and data were stored into the receiver for 48 hours. During the pH study patients were encouraged to engage in their usual activities including work, exercise, etc. with the only restrictions that no sipping of acidic beverages (eg, orange juice, cola) occurred between meals not to alter readings. They were asked to press the event marker button on the receiver whenever they experienced their symptoms. Furthermore, patients kept diaries documenting timing of food intake, change of posture and occurrence of symptoms according to their watch, which was synchronized with the Bravo receiver. After 48 hours of pH recording, the patients returned to the hospital and gave receivers and diaries back. Recorded data were downloaded onto a computer using a commercial software (Polygram Net, Medtronic, Denmark), and the second 48-hour recording began in 13 and 23 patients of the Italian and the British center respectively.

Outcome Evaluation

Patient’s outcome data were obtained by reviewing hospital files with regards to consultations occurring after wireless pH monitoring and/or by telephone interview with a standardized questionnaire.\textsuperscript{28} Heartburn was classified as same/worse, better or absent. Outcome was considered positive if heartburn was better or absent. Regarding treatment for GERD, information was sought on (1) presence of PPI treatment and whether the PPI dose was increased or not compared with pre pH test and (2) occurrence of fundoplication.

Statistical Methods

pH variables were analysed in 24-hour periods in order to determine if tracings were indicative of the diagnosis of NERD; ie, either AET > 5% or normal AET and SI ≥ 50% for heartburn, provided patients had at least 3 heartburn episodes/day; in case heartburn episodes/24-hour were less than three, the SI was not evaluated and classification of patients as FH or not was based solely on AET data. Patients with less than 3 symptomatic episodes were excluded form SI evaluation because symptom/reflux association indexes are known not to be reliable when the number of episodes is very low.\textsuperscript{32} Based on these criteria patients were reclassified as FH or NERD at the end of the test. Furthermore, percentage time at pH < 4 in upright and supine position were calculated in tracings of patients who had a change in their diagnosis because of increased AET. Only days with at least 21-hour recordings were included.

Presence of other esophageal symptoms (acid regurgitation and chest pain) was also considered in order to better characterize patients. SI was calculated separately for regurgitation and chest pain only if at least three episodes/day occurred. Moreover a global SI (heartburn, regurgitation, and chest pain) was calculated for each patient. The number of symptomatic events (both heartburn and the other symptoms) for every single 24-hour of recording and for the whole monitoring were also collected.

Data were expressed as median (first and third quartile). Non parametric statistics and chi-square test were used when appropriate. ANOVA followed by Fisher’s test was used to compare variability of occurrence of symptoms among the 4 days.

Results

All patients (N = 50) completed at least 48 hours of the prolonged wireless pH-study. Out of the 36 patients prolonging monitoring beyond 48 hours, capsule detachment occurred between 48 and 72 hours in 9 patients, and between 72 and 96 hours in 12 patients. Thus, 96-hour recordings were available in 15 of the 50 patients. All patients reported normal activities of daily living during the wireless pH-study.

Impact of Prolonged pH-monitoring on Diagnosis

Gastro-esophageal reflux

Fifteen (3 male, 55 years [range, 46-59 years]) out of the 50 patients showed a pathological AET after the first 24 hours of monitoring (10 in the second day, 2 in the third day, and 3 in the last day of recording), which changed their diagnosis from FH to NERD, whereas in the remaining 35 the diagnosis of FH was

Vol. 21, No. 2   April, 2015 (265-272)
confirmed. In particular, the number of patients with FH diminished from 50 to 40 after 48 hours, to 38 after 72 hours, and to 35 after 96 hours. Within the patients with change in diagnosis, one had an AET > 5% in all three days, 1 in 2 consecutive days and all the other 13 patients had a single-day pathological AET. AET over the 4 days in NERD patients is described in Figure.

**Symptom reflux association**

Twenty-three out of 50 patients experienced three or more episodes of heartburn during at least one day of monitoring. In one patient only SI for heartburn became positive after the first day of recording, changing the diagnosis to NERD, but this had no added value after the AET analysis, as it occurred together with pathological AET.

No patients had a positive SI for acid regurgitation or chest pain during the first 24 hours. Only one patient reported a positive SI for regurgitation during the second and the third day of monitoring, suggesting the diagnosis of NERD.

Calculation of a cumulative SI for all 3 symptoms together proved to be of no additional benefit since it was positive only in one patient, the same one with positive SI for acid regurgitation.

**Characteristics of Functional Heartburn Versus Non-erosive Reflux Disease**

**Proton pump inhibitor refractoriness**

All patients were refractory to a PPI treatment of at least 12 weeks. In the FH group we were able to collect the information on the PPI dose pre pH test in 28 patients and 18 of them (64%) underwent a double dose. In the NERD group the information was collected in 10 patients and 7 (70%) underwent a double dose ($P = 0.743$ vs FH). In the FH group we retrieved the information of partial or no response to PPIs in 29 patients and 18 (62%) had a partial response. In the NERD group the information was retrieved in 13 patients and 12 (92%) had a partial response ($P < 0.05$ vs FH).

**pH monitoring and clinical features**

AET description of the 2 groups is detailed in Table 1. AET was not different between FH and NERD patients in the first day, whereas it was so in the remaining three days ($P < 0.01$).

Considering daily differences within each group, in the 8 FH and in the 7 NERD patients having 96 hours of recordings, AET was higher in day 2 compared to day 4 in FH ($P < 0.05$) whereas it was higher in day 4 compared to day 1 in NERD ($P < 0.05$).

Upright and supine values were not available for 2 of the 15 NERD patients. Considering the days with pathological 24-hour AET, 3 out of the remaining 13 patients showed pathological AET both in upright and supine periods, whereas 7 patients had increased upright esophageal acid exposure and normal supine AET ($10.7 \pm 7.5-13.8\text{h}$ and $0 \pm 0-0.2\text{h}$); isolated pathological supine AET was seen in the remaining 3 patients.

Number of HB episodes in the first day were not different between FH and NERD (Table 2) and so was during the whole monitoring, 3 (0-6) vs 4 (1-12). Moreover, within each group, frequency of heartburn was similar in the 4 days of monitoring (Table 2).

Similarly to HB, no differences were seen between FH and NERD in occurrence of regurgitation or chest pain over the

**Table 1. Acid Exposure Time, Expressed as % Time at pH < 4, in Functional Heartburn and Non-erosive Reflux Disease Patients**

|           | Day 1 (35) | Day 2 (35) | Day 3 (17) | Day 4 (8) |
|-----------|------------|------------|------------|-----------|
| FH        | 0.9; 0.5-2.8 | 1.6; 0.6-4.1 | 1.4; 0.3-2.4 | 0.9; 0.3-1.4 |
| NERD      | Day 1 (15) | Day 2 (15) | Day 3 (10) | Day 4 (7) |
| Total     | 2; 1.3-2.6 | 6.1; 3.5-8.9 | 4.2; 2.7-6.1 | 6.1; 4-9.8 |
| Upright   | 2.9; 1.8-3.5 | 6.9; 3.2-8.8 | 5.9; 2.9-10.7 | 7.2; 4.4-12.4 |
| Supine    | 0; 0-0.3 | 0; 0-0.9 | 0; 0.3-1.5 | 2.9; 1.3-13.7 |

*Data available only for 13 patients; $P < 0.01$ vs FH patients.

FH, functional heartburn; NERD, non-erosive reflux disease.

Data are expressed as median (interquartile range [IQR]; Number of patients in brackets.
Table 2. Number of Symptomatic Episodes in Functional Heartburn and Non-erosive Reflux Disease Patients

|       | Heartburn | Regurgitation | Chest pain |
|-------|-----------|--------------|------------|
| FH (35) |           |              |            |
| Day 1  | 1 (0-4)   | 0 (0-2)      | 0 (0-2)    |
| Day 2  | 1 (0-4)   | 0 (0-2)      | 0 (0-2)    |
| Day 3  | 0 (0-5)   | 0 (0-2)      | 0 (0-3)    |
| Day 4  | 0 (0-1)   | 0 (0-0)      | 0 (0-2)    |
| NERD (15) |          |              |            |
| Day 1  | 1 (0-4)   | 0 (0-0)      | 0 (0-1)    |
| Day 2  | 1 (0-4)   | 0 (0-2)      | 0 (0-0)    |
| Day 3  | 1 (0-4)   | 1 (0-7)      | 0 (0-1)    |
| Day 4  | 1 (0-4)   | 0 (0-3)      | 0 (0-1)    |

FH, function.
Data are expressed as median (interquartile range [IQR]). Number of patients in brackets.

Discussion

Our main findings were that one third of endoscopy negative heartburn patients refractory to PPIs, diagnosed as FH during the first 24 hours of pH monitoring were actually patients with GERD as detected during a more prolonged pH recording. This observation, which underlies limitations of 24-hour pH studies, should be considered clinically relevant because, whereas the diagnosis of FH should induce the physician to stop prescribing unnecessary PPI treatment, the diagnosis of GERD related heartburn should both support a trial of high dose prolonged pharmacological treatment of GERD and open to the therapeutic option of antireflux surgery.

As a matter of fact outcome data, obtained in the majority of patients, showed that physicians increased the dose of PPIs or indicated antireflux surgery more frequently in our NERD than in our FH patients. However, they also showed that physicians often do not pay attention to a negative pH study continuing to prescribe PPIs and even sending their patients to antireflux surgery, as already observed in our previous study.28

Furthermore, our findings suggest that previous studies looking at morphological and/or pathophysiological differences between FH and NERD, diagnosed with 24-hour pH/pH-impedance studies,10-14 may have been inaccurate. Structural characteristics, which are devoid of day to day fluctuations, ie, histopathological parameters or basal impedance, could be the best variables in order to diagnose FH, but their diagnostic performance needs to be evaluated against a gold standard.48-96 hours pH monitoring could serve for this purpose. AET is known to vary from day to day,15-17 thus opening to the possibility of a false negative result for GERD. Prolonging pH monitoring to 48 hours allowed us to detect 10 patients with pathological AET and further prolonging the test to 96 hours allowed to detect 5 more patients, in line with previous data showing that 96 hours is better than 48 hours in the diagnosis of GERD.28-30 AET was the most important variable in order to diagnose GERD after the first 24 hours of pH monitoring. In one patient only the heartburn/reflux association changed from negative to positive in the second day of recording and it was accompanied by pathological AET.

We also looked at the possibility that inclusion of other symptoms, such as regurgitation or chest pain, in the calculation of symptom/reflux association increases the power of symptom/reflux association indexes for exclusion of FH. However, it was not the case because regurgitation and chest pain were quite in-
frequent in this patients’ population and, in one patient only, addition of regurgitation has suggested to exclude the diagnosis of FH.

When looking at the distribution of acid reflux in the 24 hours in our patients with pathological AET, 76% had increased reflux during daytime in line with the evidence that NERD patients are mainly postprandial upright refluxers.33

It could be argued that during the day of capsule positioning patients may have eaten and moved less than during the following days because of the discomfort of the endoscopy and sedation. We do not think this to be the case for several reasons. First, the literature looking at trends of AET during the days of prolonged pH monitoring has generally reported similar AETs, more rarely higher or lower AET in the first day of recording compared to the others. As a matter of fact our data have shown that AET in day 2 was not significantly higher than AET in day 1 both in patients with normal and in those with pathological reflux. Furthermore 13 out of the 15 patients who were diagnosed as NERD had pathological AET during one of the days only. Finally, the mild i.v. sedation we used is known to allow almost full psychomotor recovery 30 minutes after upper endoscopy.39

Another information we looked for in our 50 patients was a difference in demographics, clinical presentation, characteristics of reflux or symptoms during the first 24 hours of pH monitoring between those that remained FH and those who were diagnosed as NERD. We could not find any significant differences although numbers were relatively small. The retrospective nature of our study did not allow us to collect information on presence of comorbidities, such as functional gastrointestinal disorders, in a sufficient number of patients or whether the first day of recording was different from everyday life concerning the symptom profile. Our technique devoid of impedance monitoring could not differentiate patients with heartburn related to weakly acidic reflux, who are erroneously diagnosed as FH during 24 hours pH monitoring, as shown by a big series of endoscopy negative patients with typical GERD symptoms.9 We do not know if these patients partially overlap with those showing pathological AET after the first day of wireless pH monitoring and also if 48 hours catheter based pH/impedance monitoring is feasible and improves the diagnosis of GERD. Further studies could answer these questions.

Results on outcome should be cautiously considered due to the retrospective nature of our study. Nevertheless it showed a positive impact of the pH test on management and outcome of patients classified as NERD. Not all of the NERD patients, however, had satisfactory outcome, although they were all on antireflux treatment. We cannot exclude that our pH test was falsely positive. Presence of GERD was assessed with the “worst day” analysis which is more sensitive than the “whole period analysis,” but it is presumably less specific.40 Regarding patients with FH, it might seem weird that 45% of them had a positive outcome, but reassurance after a negative diagnostic test, ie, upper gastrointestinal endoscopy, has already been shown to diminish anxiety, improve quality of life and decrease consultations in patients with dyspepsia.41,42 Furthermore, placebo effect of drugs is frequent in functional disorders.43 As a final consideration, we need to point out that our paper relates to a “real world” situation, involving those patients that are referred to our three Centers for wireless pH monitoring, and not to a clinical trial. Compliance to treatment was not recorded and in one-third of our patients refractoriness was based in a standard dose PPI. It is thus possible that some of our patients were PPI responders at higher doses, as suggested by our follow up data and that wireless pH monitoring was not indicated. Nevertheless in this “real world” scenario wireless pH monitoring was useful for patients’ management. A prospective study with a bigger cohort of patients is warranted.

In conclusion, our data underlie the moderate reliability of 24 hours reflux studies in the diagnosis of FH. They suggest usefulness of prolonged wireless pH monitoring in selected cases of clinical practice where there is a strong clinical suspicion of GERD, but normal 24 hours reflux monitoring.

Acknowledgements

We thank Dr Terry Wong and Dr Angela Anggiansah, Physiology Lab, Guys and St. Thomas’ NHS Foundation Trust, London, UK for allowing Dr Rami Sweis to collect and use data of prolonged wireless pH monitoring performed in their Laboratory.

References

1. Dent J, El-Serag HB, Wallander MA, Johansson S. Epidemiology of gastro-oesophageal reflux disease: a systematic review. Gut 2005; 54:710-717.
2. Bruley Des Varannes S, Marek L, Humeau B, Lecasble M, Colin R. Gastroesophageal reflux disease in primary care. Prevalence, epidemiology and quality of life of patients. Gastroenterol Clin Biol 2006;30:364-370.
3. Klauser AG, Schindlbeck NF, Müller-Lissner SA. Symptoms in gastro-oesophageal reflux disease. Lancet 1990;335:205-208.
4. Lind T, Haveland T, Carlsson R, et al. Heartburn without oesopha-
Functional Heartburn and Wireless pH Monitoring
prolonged pH monitoring using the Bravo system in Mexico. Am J Gastroenterol 2005;100:2382-2386.

38. des Varannes SB, Mion F, Ducrotté P, et al. Simultaneous recordings of oesophageal acid exposure with conventional pH monitoring and a wireless system (Bravo). Gut 2005;54:1682-1686.

39. Willey J, Vargo JJ, Connor JT, Dumot JA, Conwell DL, Zuccaro G. Quantitative assessment of psychomotor recovery after sedation and analgesia for outpatient EGD. Gastrointest Endosc 2002;56:810-816.

40. Grigolon A, Bravi I, Duca P, Pugliese D, Penagini R. Prolonged wireless pH monitoring: importance of how to analyse oesophageal acid exposure. Scand J Gastroenterol 2010;45:1133-1134.

41. Hungin AP, Thomas PR, Bramble MG, et al. What happens to patients following open access gastroscopy? An outcome study from general practice. Br J Gen Pract 1994;44:519-521.

42. Lucock MP, Morley S, White C, Peake MD. Responses of consecutive patients to reassurance after gastroscopy: results of self administered questionnaire survey. Bmj 1997;315:572-575.

43. Shah E, Pimentel M. Placebo effect in clinical trial design for irritable bowel syndrome. J Neurogastroenterol Motil 2014;20:163-170.

44. Cantù P, Savojardo D, Carmagnola S, Penagini R. Impact of referral for gastro-oesophageal reflux disease on the workload of an academic Gastroenterology Unit. Dig Liver Dis 2005;37:735-740.

45. Dickman R, Boaz M, Aizic S, Beniashvili Z, Fass R, Niv Y. Comparison of clinical characteristics of patients with gastroesophageal reflux disease who failed proton pump inhibitor therapy versus those who fully responded. J Neurogastroenterol Motil 2011;17:387-394.