

Correlation between the different pH-metry scores in gastroesophageal reflux disease in children

Vasile Valeriu Lupu (MD, PhD)^{a,b}, Ancuta Ignat (MD)^{a,*}, Gabriela Paduraru (MD, PhD)^{a,b}, Anamaria Ciubara (MD, PhD)^c, Mihaela Moscalu (PhD)^d, Cristina Oana Marginean (MD, PhD)^e, Marin Burlea (MD, PhD)^{a,b}

Abstract

The 24-hour esophageal pH-metry is the most widely used method to diagnose the gastroesophageal reflux disease (GERD). The study compares the different scores obtained during the 24-hour esophageal pH-metry. A retrospective study over 5 years including 234 children (1 month and 18 years old) admitted in a pediatric gastroenterology regional center in Northeast Romania, with suspicion of GERD. They underwent 24-hour esophageal pH-metry, and the scores obtained (Boix-Ochoa, DeMeester, Johnson-DeMeester) were compared. Out of the 234 children, 172 (73.50%) had positive Boix-Ochoa score and 62 (26.50%) had normal Boix-Ochoa score (<11.99). Based on the DeMeester score, 149 children (63.68%) were positive and 85 (36.32%) were negative. The correlation of the Demeester score with the Boix-Ochoa score was very high ($r=0.978$, $P < 0.01$, 95% confidence interval). Considering the Johnson-DeMeester score, 120 cases (51.28%) had GERD and 114 (48.72%) did not. The correlation of the Johnson-DeMeester score with the Boix-Ochoa score was still high ($r=0.94$, $P < 0.01$, 95% 95% confidence interval). As considered until now, the Boix-Ochoa score is the most accurate score to be used in pediatrics for the diagnosis of GERD. The use of the different scores—Boix-Ochoa, DeMeester, Johnson-DeMeester—showed a high sensitivity and specificity of the pH-metry measurements applied to the study lot, but the last score has a higher risk of false-negative results.

Abbreviations: GERD = gastroesophageal reflux disease, Se = sensitivity, Sp = specificity.

Keywords: Boix-Ochoa score, child, Demeester score, esophageal pH-metry, gastroesophageal reflux disease, Johnson-DeMeester score, sensitivity and specificity

1. Introduction

Considered the most accessible method in the diagnosis of gastroesophageal reflux, the pH-metry is based on the principle that the passage of acid gastric contents into the esophagus during the reflux causes a decrease in the esophageal pH. The method consists of recording the pH in the lower esophagus over a period of time (in our case 24 hours), with an electrode placed 5 cm above the cardia, which is connected to a portable, battery-powered, pH meter. The pH can be written on paper or

transferred to a computer, which analyzes various required parameters, resulting in a graphical representation of esophageal pH curves.^[1]

The test is considered positive if the pH falls below 4 for a period longer than 5 seconds.^[2] In interpreting the results, the following parameters were considered: total number of reflux episodes (normal <2 episodes per hour on average), number of reflux episodes lasting more than 5 minutes (normal <8 episodes), duration of the longest reflux episode (in minutes), reflux index (RI)=ratio between the total number of reflux episodes and recording time (normal <4), Euler score= $x+4y$ (where x is the number of reflux episodes with pH <4 longer than 1 minute and y is the number of episodes with pH <4 longer than 5 minutes).^[2]

After the computer interpretation of data recorded, we can obtain 3 scores: Boix-Ochoa (n <11.99), DeMeester (n <14.72), and Johnson-DeMeester (n <22). The scores were calculated in the same time by the program.

With small-diameter probes, the method can be used also to premature babies, being well-tolerated and easy to perform. Depending on the height of the child, the electrode position is determined using age tables or by radiological control of the electrode's position.^[3]

The literature recommends for pediatric age the use of Boix-Ochoa score. Being able to calculate other scores also (DeMeester and Johnson-DeMeester), we proposed to study possible correlations and to determine their sensitivity and specificity.

The presence of typical findings of reflux esophagitis on endoscopy is a diagnose of gastroesophageal reflux disease

Editor: M Jose Carbonero Cels.

The authors have no funding and conflicts of interest to disclose.

^aPediatrics Department, University of Medicine and Pharmacy "Gr. T. Popa,

^bVth Pediatrics Clinic, "St. Mary" Children Emergency Hospital, ^cPsychiatry Department, ^dMedical Informatics and Biostatistics Department, University of Medicine and Pharmacy, "Gr. T. Popa", Iasi, Romania, ^ePediatrics Department, University of Medicine and Pharmacy, Targu Mures, Romania.

* Correspondence: Dr Ancuta Ignat, str. Universitatii, nr. 16, Iasi 700115, Romania (e-mail: anca_ign@yahoo.com).

Copyright © 2016 the Author(s). Published by Wolters Kluwer Health, Inc. All rights reserved.

This is an open access article distributed under the Creative Commons Attribution-NoDerivatives License 4.0, which allows for redistribution, commercial and non-commercial, as long as it is passed along unchanged and in whole, with credit to the author.

Medicine (2016) 95:26(e3804)

Received: 8 October 2015 / Received in final form: 17 March 2016 / Accepted: 26 April 2016

<http://dx.doi.org/10.1097/MD.0000000000003804>

(GERD) with a specificity of 90% to 95%.^[4] The endoscopic findings of reflux esophagitis in the lower esophagus were classified according to the Los Angeles classification system. Esophagitis was graded by endoscopy: grade A, 1 (or more) mucosal break no longer than 5 mm, that does not extend between the tops of 2 mucosal folds; grade B, 1 (or more) mucosal break more than 5 mm long that does not extend between the tops of 2 mucosal folds; grade C, 1 (or more) mucosal break that is continuous between the tops of 2 or more mucosal, folds but which involves less than 75% of the circumference; grade D, 1 (or more) mucosal break which involves at least 75% of the esophageal circumference.^[5]

2. Methods

We performed a retrospective study on 234 children of both sexes, 1 month and 18 years old, both from rural and urban area, admitted in a pediatric gastroenterology regional center in Northeast Romania, over a period of 5 years. They underwent 24-hour esophageal pH-metry. The indications for 24-h pH monitoring involved a diagnostic uncertainty. The patients had atypical symptoms, especially respiratory, unresponsive to usual therapy, such as cough, hoarseness, sore throat, atypical chest pain, and asthma. The cases which did not match with the imposed criteria (the children with a diagnostic certainty) were excluded from the study. Informed consent was obtained from all patients/caregivers, and the "St. Mary" Children Emergency Hospital Ethics Committee's approval was obtained for publishing this study.

To determine the esophageal pH, we used Medtronic Digitrapper pH 100, SN 37660, with Polygram Net TM pH Testing Application and Zinetics 24 multi-use and ComforTec by Sandhill catheters. All measurements were made over a period of 24 hours.

The device is calibrated in 2 solutions with pH 1 and 7 before each use. Before the procedure the patient should have not ingested food and drink at least 6 hours for children over 1-year-old, at least 3 hours for infants.^[6] Antacid therapy should be discontinued at least 6 hours before, the anti-H2 receptor 3 days before, the proton pump inhibitor 7 days before, and the prokinetics 48 hours before.^[7]

The child is placed in the left lateral decubitus (for infants and young children), the examiner is to the right side of the patient, and in a seated position for children over 5 to 6 years old. The electrode is inserted, lubricated before, nasal up to 5 cm above the cardia. Then the electrode is connected to the pH meter, the function and operation of the device are explained to the caregiver/patient, and then, the recording is started. Each patient or caregiver is trained to record in a table any symptoms occurred, the time and body position (supine, standing), and, at the same time, using the button provided on the device.

The pH-metry was conducted and the results were interpreted by a pediatrician, who is a subspecialist in pediatric gastroenterology.

The SPSS 20 was used for the statistical data processing. For the correlation analysis, the Pearson parametric correlation was used and the correlation coefficients were calculated for a confidence interval (CI) of 95%. The Boix-Ochoa score, as a gold standard, was used to determine the positive and negative predictive value of DeMeester and Johnson-DeMeester scores. The literature recommends for pediatric age the use of Boix-Ochoa score.

3. Results

3.1. PH-metry interpretation by the Boix-Ochoa score ($n < 11.99$), DeMeester score ($n < 14.72$), and Johnson-DeMeester score ($n < 22$)

In the past 5 years, out of 234 children (1 month and 18 years old) of both sexes, both from rural and urban area, 172 (73.50%) had a positive Boix-Ochoa score and 62 (26.50%) had a normal Boix-Ochoa score (< 11.99). As regards to the DeMeester score, in 149 children (63.68%) it was positive and in 85 (36.32%) it was normal (< 14.72). According to the Johnson-DeMeester score, gastroesophageal reflux was associated in 120 cases (51.28%), and it was not associated in 114 cases (48.72%) (Table 1). The scores were calculated in the same time by the program. There were not any adverse events registered after pH-metry.

In a batch of 72 patients, GERD was also confirmed, based on upper gastrointestinal endoscopy. *Helicobacter pylori* was found in 19 (26.39%) of the 72 patients, and in 53 (73.61%) patients it was absent, and in the evolution of the GERD, no particular clinical or evolutional aspects related to the presence or absence of the *Helicobacter pylori* infection were noted.

3.2. Correlation of the DeMeester score versus the Boix-Ochoa score

The correlation of the DeMeester score with the Boix-Ochoa score is very high ($r=0.978$, $P < 0.01$, 95% CI) (Table 2).

The results show good values of sensitivity (Se) and specificity (Sp) of the DeMeester test. When the DeMeester score was applied, 86.63% of the cases confirmed with GERD by the Boix-Ochoa score were also confirmed by the DeMeester score, and only 13.37% were false-negative (Table 3). It is also ascertained that the DeMeester score did not show any false-positive case; therefore all negative cases were correctly diagnosed (Table 4).

In conclusion, when applying the DeMeester test, we take the risk of potential false-negative results ($Sp > Se$, $Sp = 98.4\%$).

3.3. Correlation between the Johnson-DeMeester score versus the Boix-Ochoa score

The correlation of the Johnson-DeMeester score with the Boix-Ochoa score is very high ($r=0.94$, $P < 0.01$, 95% CI), but it is slightly lower in comparison with the DeMeester score (Table 5).

Table 1

Ph-metry interpretation by the Boix-Ochoa score, DeMeester score, and Johnson-DeMeester score.

Boix-Ochoa score positive	Boix-Ochoa score normal	DeMeester score positive	DeMeester score normal	Johnson-DeMeester score positive	Johnson-DeMeester score normal
172 (73.50%)	62 (26.50%)	149 (63.68%)	85 (36.32%)	120 (51.28%)	114 (48.72%)
Total 234					

Table 2

Results of the Pearson correlation test between DeMeester score and Boix-Ochoa score.

Pearson	r (correlation coefficient) (95% confidence interval)	P
DeMeester score vs Boix-Ochoa score	0.9781	0.00

Table 3

Congruence of pH-metry versus DeMeester score.

	DeMeester negative		DeMeester positive		
	No. of cases	%	No. of cases	%	
pH-metry	+	23	13.37%	149	86.63%
	-	62	100%	0	
Total		234			

Table 4

DeMeester score sensitivity and specificity.

DeMeester score	
Sensitivity, %	86.62
Specificity, %	98.43
Positive predictive value, %	99.33
Negative predictive value, %	73.25
Accuracy, %	90.21

Table 5

Results of the Pearson correlation test between the DeMeester score and the Boix-Ochoa score.

Pearson	r (correlation coefficient) (95% confidence interval)	P
Johnson-DeMeester score vs Boix-Ochoa score	0.9409	0.000

The results show very good values of the specificity of the Johnson-DeMeester test, which proves the summing of the false-negative results by considering the results of this table (Table 6). Technically, this test increases the risk of the presence of gastroesophageal disease, which may, however, not be detected after the Johnson-DeMeester test.

In comparison with the DeMeester test, the Johnson-DeMeester test has a higher risk of false-negative results (Table 7).

It should be noted that the motivation of the pH-metric investigation was rather the associated pathology than the

Table 7

Johnson-DeMeester score sensitivity and specificity.

Johnson-DeMeester score	
Sensitivity, %	69.76
Specificity, %	98.41
Positive predictive value, %	99.17
Negative predictive value, %	54.38
Accuracy, %	77.44

digestive symptomatology. Our results assessed, based on Boix-Ochoa, Demeester, and Johnson-DeMeester scores, show a high degree of specificity and sensitivity when applied to our patients.

We do not have indeterminate results, missing data, and outliers of the index tests.

4. Discussion

The North American Society of Pediatric Gastroenterology, Hepatology, and Nutrition, and the European Society for Pediatric Gastroenterology, Hepatology, and Nutrition guidelines on pediatric GER established that multichannel intraluminal impedance-pH monitoring detects acid, weak acid, and nonacid reflux episodes, and it is superior to pH monitoring alone for evaluation of the temporal relation between symptoms and GER.^[8] This technique is limited by its high cost and the lack of normative data of multichannel intraluminal impedance-pH in the pediatric population.^[9] However, as long as there is no effective medical therapy for weak acid and nonacid reflux, the clinical relevance of measuring these types of reflux remains debatable.^[10]

Nowadays, pH-metry is the most widely used method of diagnosis and assessment of patients with typical or atypical GERD symptoms. It is also the method most widely used in patients with reflux episodes under treatment with acid-suppressing medication, when symptoms persist. The great advantage is the possibility of outpatient monitoring, and sensitivity and specificity are high: 90% sensitivity and 100% specificity in patients with typical symptoms.^[11]

The sensitivity and specificity of a method depend only on the test qualities (and potentially on the examiner), whereas the predictive values do not provide information on the test itself, but on the situation after the exploration, being related to the clinical context (pretest probability), and also the native characteristics of the method (sensitivity and specificity). Furthermore, the more sensitive the test is, the better the negative predictive value is (the degree of certainty that a patient with a normal test does not have the disease is higher), and the more specific the test is, the better the positive predictive value is; therefore a patient with an abnormal test will definitely have the disease. Accuracy shows us how close we are to the actual value of the test. Accuracy is the extent to which the measurement or the estimation based on certain measurement presents the actual value of the measured characteristic.

When pH evaluation is performed over a prolonged period, it presents high sensitivity and specificity indices.^[11] The Boix-Ochoa methodology is considered to be the most appropriate for application to the pediatric age group.^[12]

The DeMeester score has its limitations, with reported sensitivities ranging from 60% to 100%,^[13] through to indexes as low as 28%.^[14] The inconstant sensibility of the examination made its methodology doubtful, and then some points were considered if they were normal: gastroparesis or functional

Table 6

Congruence of the results of the pH-metry versus Johnson-Demeester score.

	Johnson-Demeester negative		Johnson-Demeester positive		
	No. of cases	%	No. of cases	%	
pH-metry	+	52	30.23%	120	69.77%
	-	62	100.00%	0	
Total		234			

heartburn, the presence of nonacid reflux, the possibility that pH probe might have missed distal acid reflux, noxious effect of the nasal catheter could have limited both eating and activity and resulted in a false-negative test, and the influence of electrode location in the result of esophageal prolonged pH monitoring.^[15]

The sensitivity of Johnson-Demeester score for the typical symptom heartburn is 68% and its specificity is 63%.^[16] In our study, we obtained a sensitivity similar to that of literature—69.76% and a higher specificity of 98.41%, but we applied pH-metry only for atypical symptoms. Since its description in 1974, the 24-hour esophageal pH composite score has proved to be a durable standard to identify and quantify GERD by measuring esophageal acid exposure.^[17]

With an accurate score and more experience in pH monitoring, we would be able to identify children at risk to develop severe complications, control them, and aid in the selection of those patients requiring surgical intervention.

The authors of a study stated that pH monitoring has false-positive results if using conventional parameters, which means stringent methodological criteria should be applied.^[18]

Study limits are dictated by monitoring body position in young children in a hospital ward (not in their normal environment at home), which presents considerable practical difficulties, as few children maintain any position for very long. Also, the pH monitoring alone, can not diagnose alkaline reflux. Another limitation would be that we compared the scores only for reflux with atypical symptoms.

5. Conclusions

As considered until now, the Box-Ochoa score is the most accurate score to be used in pediatrics for the GERD diagnose. The use of the different scores—Box-Ochoa, Demister, and Johnson-Demister—showed a high sensitivity and specificity of the pH-metric measurements applied to the study lot. Regarding the Demister score, we obtained a sensitivity of 86.62% and a specificity of 98.43%. The Johnson-Demister score had a higher risk of false-negative results and we found a sensitivity of 69.76% and a specificity of 98.41%.

Acknowledgments

We would like to thank the Endoscopy Department's staff, and the Vth and IIInd Pediatric Clinic staff at "St. Mary" Children's Emergency Hospital for their help.

References

- [1] Christopher G, Streets MRCS, Tom R, DeMeester . Ambulatory 24-hour esophageal pH monitoring: why, when, and what to do. *J Clin Gastroenterol* 2003;37:14–22.
- [2] Chiou E, Rosen R, Jiang H, et al. Diagnosis of supra-esophageal gastric reflux: correlation of oropharyngeal pH with esophageal impedance monitoring for gastroesophageal reflux. *Neurogastroenterol Motil* 2011;23:717-e326.
- [3] Walker WA, Kleiman RE, Sherman PM, et al. *Pediatric Gastrointestinal Disease*. IVth ed.2004;Hamilton:BC Decker Inc, 384–399.
- [4] Moayyedi P, Tally N. Gastro-oesophageal reflux disease. *Lancet* 2006;367:2086–100.
- [5] Lundell LR, Dent J, Bennett JR, et al. Endoscopic assessment of oesophagitis: clinical and functional correlates and further validation of the Los Angeles classification. *Gut* 1999;45:172–80.
- [6] Jamieson JR, Stein CM, De Meester TR, et al. Ambulatory 24-h esophageal pH monitoring: normal values, optimal thresholds, specificity, sensitivity, and reproducibility. *Am J Gastroenterol* 1992;87:171–5.
- [7] Kleiman RE, Goulet OJ, Mieli-Vergani G, et al. Walker's pediatric gastrointestinal disease. 5th ed. BC Decker Inc, 2008;(2), 1393–1400.
- [8] Vandenplas Y, Rudolph CD, Di Lorenzo C, et al. Pediatric gastroesophageal reflux clinical practice guidelines: joint recommendations of the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN) and the European Society for Pediatric Gastroenterology, Hepatology and Nutrition (ESPGHAN). *J Pediatr Gastroenterol Nutr* 2009;49:498–547.
- [9] Shin MS. Esophageal pH and combined impedance-pH monitoring in children. *Pediatr Gastroenterol Hepatol Nutr* 2014;17:13–22.
- [10] Wenzl TG1 , Benninga MA, Loots CM, et al. Indications, methodology, and interpretation of combined esophageal impedance-pH monitoring in children: ESPGHAN EURO-PIG standard protocol. *J Pediatr Gastroenterol Nutr* 2012;55:230–4.
- [11] Jamieson JR, Stein HJ, DeMeester TR, et al. Ambulatory 24-h esophageal pH monitoring: normal values, optimal thresholds, specificity, sensitivity and reproducibility. *Am J Gastroenterol* 1992;87:1102–11.
- [12] Boix-Ochoa J, Lafuente JM, Gil-Vernet JM. Twenty-four hour esophageal pH monitoring in gastroesophageal reflux. *J Pediatr Surg* 1980;15:74–8.
- [13] Jamieson JR, Stein HJ, DeMeester TR. Ambulatory 24-h esophageal pH monitoring: normal values, optimal thresholds, specificity, sensitivity, and reproducibility. *Am J Gastroenterol* 1992;87:1102–11.
- [14] Shay S, Richter J. Direct comparison of impedance, manometry, and pH probe in detecting reflux before and after a meal. *Dig Dis Sci* 2005;50:584–90.
- [15] Valter Nilton Felix , Ioshiaki Yogi , Daniel Senday , et al. Discussing the influence of electrode location in the result of esophageal prolonged pH monitoring. *BMC Gastroenterol* 2014;14:64.
- [16] Charbel S, Khandwala F, Vaezi MF. The role of esophageal pH monitoring in symptomatic patients on PPI therapy. *Am J Gastroenterol* 2005;100:283–9.
- [17] Tefera L, Fein M, Ritter MP, et al. Can the combination of symptoms and endoscopy confirm the presence of gastroesophageal reflux disease? *Am Surg* 1997;63:933–6.
- [18] Christopher G Streets, Tom R, DeMeester . Ambulatory 24-hour esophageal pH monitoring. *J Clin Gastroenterol* 2003;37:14–22.