Functional Heartburn is Gastroesophageal Reflux

Michael Levin*

Department of Pediatric Radiology of the 1-st State Hospital, Minsk, Belarus, Dorot. Medical Center for Rehabilitation and Geriatrics, Amnon veTamar, Netanya, Israel

ABSTRACT

Most articles on the diagnosis and treatment of patients with GERD are based on recommendations from conferences participants that voted using the Delphi method. However, this method is intended for the organization of production and social planning, not for scientific analysis. This article analyzes two review articles on the differential diagnosis of GERD and functional heartburn. These diseases are not distinguishable by their clinical presentation. Differential diagnosis relies only on pH monitoring in the esophagus. In the presence of weakly acidic reflux (pH < 4 < 6% of the time out of 24 hours of recording), reflux is considered physiological, because of which GERD is excluded. These articles were published 5 years apart and the authors cite different conferences. As a result of the repeated voting, the definition of GERD was changed, and the diagnosis was now based only on pH monitoring and baseline impedance, if endoscopy did not reveal pathology. An analysis of the literature shows that pH monitoring detects only severe forms of GERD. In more than 30% of patients, the diagnosis of GERD is not confirmed, and they are unreasonably diagnosed with functional heartburn or other functional diseases. These patients do not receive timely pathogenetic treatment, which leads to chronicity of the process and severe complications. Voting participants, without scientific grounds, consider histological studies to be of little information to present pH monitoring as the gold standard in the diagnosis of GERD, which it is not. Conclusion. An analysis of these articles showed that the decisions of various conferences on the diagnosis and treatment of gastroesophageal reflux, made using the Delphi method, are not scientific. Their recommendations often contradict known scientific facts, contradict each other and sometimes to common sense. Because of using pH monitoring as a gold standard, which it is not, about two thirds of patients with GER are diagnosed with various supposedly functional diseases and they do not receive timely pathogenic treatment, which leads to chronicity of the process and severe complications. A free discussion is needed to eliminate contradictions in modern gastroenterology.

Keywords: Gastroesophageal Reflux Disease, Functional Heartburn, Delphi Method, Montreal Consensus, Lyon Consensus, Discussion

*Corresponding Author

Dr. Michael Levin,

Department of Pediatric Radiology of the 1-st State Hospital, Minsk, Belarus, Dorot. Medical Center for Rehabilitation and Geriatrics, Amnon veTamar, Netanya, Israel, Tel: 972-538281393, E-mails: nivel70@hotmail.com; michael.levin@dorot.health.gov.il

INTRODUCTION

Medicine is a science that is developed through scientific research. Any research begins with an assumption (hypothesis), which should not contradict reliable scientific facts. If multiple use hypothesis does not reveal contradictions, and is accompanied by a clinical effect, then it moves to the stage of theory. If at least one of the reliable studies contradicts the hypothesis it must be either corrected or rejected. The concept of functional heartburn, as a disease that has nothing in common with GERD except for the symptom of heartburn, appeared because of a vote by gastroenterologists using the Delphi method.

IS THE DELPHI METHOD A SCIENTIFIC METHOD?

The main purpose of the Delphi method is to encourage the experts to settle on a mutual agreement and to establish a group consensus. Many industries and organizations may use this method for business forecasting or structural decisions, like industry predictions, government planning or financial strategies. If experts can all come to a logical agreement that could be beneficial to your organization, you can also feel confident knowing the input you're receiving is truthful. The disadvantage of using the Delphi is preventing live discussions from occurring. This suggests that the consensus that participants reach is not always the best option. In addition, responses could provide little to no value due to participants being unable to come to a consensus on important issues. From this common understanding of the Delphi method, it solves production problems and cannot serve as a method of scientific knowledge. Secondly, since the decisions of the participants are not based on scientific facts, conflict of interest can lead to harmful decisions. For example, if some of the participants in the vote received grants from a competing organization, then the consensus can lead to severe financial losses. From this analysis the Delphi method is not a scientific method. Moreover, I have not found a single study in which representatives of any science used it for scientific purposes.

To substantiate my claims, I analyzed two articles on the differential diagnosis of GERD and functional heartburn (FH), published in 2019 [1] and 2024 [2]. The first article was based on the 2006 Montreal Consensus, in which "GERD was defined as a condition that develops when the reflux of stomach contents causes troublesome symptoms and/or complications" [3]. 44 experts from 18 countries participated in the voting, including 5 gastroenterologists who regularly receive grants for publishing review articles and 44 gastroenterologists whose participation was also paid by PHARMA [4]. These data indicate that the interests of manufacturers of diagnostic equipment and/or drugs motivated the decisions of the Montreal Consensus (2006).

THE DECISIONS OF THE MONTREAL CONSENSUS (2006) CONTRADICT KNOWN SCIENTIFIC FACTS, EACH OTHER AND COMMON SENSE

Before the publication of the "DeMeester scores", repeated episodes of reflux of gastric contents into the esophagus were considered to indicate reflux disease, which excluded the possibility of physiological reflux. Therefore, the disease was called "gastroesophageal reflux" (GER), and the diagnosis of GER was based on the detection of reflux episodes during radiographic examination, the detection of reflux complications during endoscopy, and a decrease in basal pressure and shortening of the lower esophageal sphincter (LES) during manometry. It was known that the cause of acid-dependent diseases of the stomach, duodenum, and esophagus is hypersecretion of hydrochloric acid. Therefore, gastric acidity was studied, and treatment was used that suppressed the secretion of hydrochloric acid and neutralized it.

In 1974 and 1976 DeMeester et al published two articles proposing a normal range for esophageal pH monitoring. It was defined as pH < 4 for 4% of the 24 hours of monitoring 5 cm proximal to the LES. The authors examined 15 individuals who believed that they had no problems with their digestive system. Since then, this boundary has been called the "DeMeester score", and the proposed method of pH monitoring has long been considered the gold standard for diagnosing gastroesophageal reflux disease (GERD) [5,6]. Firstly, it is counterintuitive because hydrochloric acid and pepsin destroy food proteins and are therefore dangerous for the esophageal wall. Even the mucous membranes of the stomach and duodenum, which are protected from the effects of chyme do not always cope with this protection, resulting in inflammatory changes up to ulcers. It is impossible to imagine that the esophagus would not be damaged by exposure to chyme in the esophagus for 50 minutes during the day. Secondly, to determine the normal range, it was necessary to very accurately select individuals without reflux. This is the ABC of scientific research. DeMeester et al. had the opportunity to use radiological, manometric and endoscopic methods to diagnose GER, as they described in another study [7] but did not use them to select individuals for the control group. DeMeester et al, could not have been unaware that GER affects more than a third of the population and can proceed for a long time without clinical manifestations. For example, endoscopic examinations of individuals who consider themselves healthy revealed GERD in 16% among 6,683 health examinees [8]. Similar results were obtained by Stål et al, who noted that "Histologic abnormalities are poorly related to acid reflux in healthy volunteers" [9]. Shieh et al showed that after POEM, 41.9% had erosive esophagitis, but only 12% had GERD symptoms

[10]. Often, GERD hides behind non-esophageal symptoms [11]. If we consider that endoscopic examination based on visual data determines only complications of GERD, it becomes obvious that the number of patients with GERD among individuals without clinical symptoms is significantly higher than shown above. This reliable data shows that the absence of clinical symptoms does not allow us to exclude GERD. In addition, the absence of complaints in patients with reflux esophagitis can be explained by damage to sensitive nerve elements by hydrochloric acid and pepsin. Consequently, endoscopic examination without histology does not allow GERD to be ruled out. Since 2018, DeMeester and Chandrasoma have recommended a biopsy of the squamocolumnar junction that in GER shows microscopic intestinalization of metaplastic cardiac mucosa [12]. They have proven that reflux begins with acid penetration only into the intra-abdominal part of the LES. Chyme penetrates the esophagus at a later stage. Microscopic intestinalization of metaplastic cardiac mucosa allows GER to be detected in time to prevent disease progression [13]. These studies show that reflux begins before chyme penetrates the esophagus. Therefore, reflux cannot be physiological. Did DeMeester realize that by signing these articles he had effectively discredited the pH monitoring he developed as a scientifically valid diagnostic method? Studies show that pH monitoring detects only severe forms of GERD. As a result, 24-hour esophageal pH measurement has a false negative rate of 15% to 30%" [8,9,14].

Based on the above, the definition of GERD by Montreal Consensus, that "GERD was defined as a condition that develops when the reflux of stomach contents causes troublesome symptoms and/or complications" [3] is unfounded and dangerous, since it leaves the early stages of GERD, when there are no troublesome symptoms and complications, outside the scope of medical observation. So instead of the diagnosis of "GER", the diagnosis of "GERD" appeared. By promoting pH monitoring as a discriminatory diagnostic method, the authors had to invent diseases where typical symptoms of GERD are not confirmed by a study of acidity in the esophagus. Functional heartburn is one of them.

Functional heartburn

In the article by Gabbard and Vijayvargiya (2019), FH is considered from the standpoint of the 2006 Montreal Consensus as the presence of heartburn with no objective evidence of GERD. It accounts for more than half of all referrals for PPI-refractory GERD. The authors emphasize that to make this diagnosis, the results of upper endoscopy with biopsy, esophageal manometry, and esophageal pH monitoring should be normal. If esophageal mucosa appears normal, biopsy of the proximal and distal esophagus should be performed to exclude an inflammatory disease such as eosinophilic or lymphocytic esophagitis. It is believed that the pathogenesis of FH is poorly understood, but may include activation of inflammatory mediators, disruption of the integrity of the esophageal mucosa, etc. It is theoretically assumed that FH is a functional or hypersensitivity disorder of the esophagus resulting from hypersensitivity of the visceral nerves of the esophagus. Therefore, neuromodulators to reduce pain perception are the basis of treatment, increased chemical and pressure sensitivity in the esophagus, as well as both peripheral and central sensitization [1].

Any hypothesis must have at least some justifications. This hypothesis is contradicted by other statements of the authors of the article.

Firstly, according to the updated recommendations of the Porto GERD consensus group, the pH test is considered positive if the acid exposure time exceeds 6% of the testing period (1.5 hours during the day). As shown above, pH monitoring diagnoses only severe GER. Obviously, those patients with a DeMeester score < 4, and especially < 6, also suffer from reflux disease.

Secondly, numerous articles have shown that with GERD, compared to the norm, dilated intercellular spaces are determined [15,16]. This accurate and cheap method, which could be a competitor to pH monitoring, was the subject of 8 articles by a regular participant in various consensus D. Sifrim, in which he reports the low reliability of this method. In the last article, London resident Sifrim, in the role of "critical revision and supervision", evaluates the results of GERD diagnostics by Turkish doctors. In this analysis, pH monitoring appears to be the gold standard for GERD diagnostics. As a control, "patients who had typical GERD symptoms (heartburn and/or regurgitation) at least once a week were included", in whom endoscopy, HRM and pH monitoring showed no pathology [17]. Comparing less severe patients, whom they groundlessly considered a control, with more severe patients, they obtained an insignificant difference (P=0.02) and for the eighth time in a row Sifrim insists on the false conclusion that this method cannot be used in everyday life. As in all previous articles, in addition to regular income for consultations and lectures, he received a grant (Ege University Scientific Research Project Coordination Unit (Project ID: TGA-2021-22732).

Five years later, from the article by Davis and Gyawali, which popularizes other recommendations of the Lyon consensus 2.0, we learn that "Mucosal damage from reflux can lead to dilated intracellular spaces, but this requires advanced techniques such as electron microscopy for optimal characterization" [2]. It follows that (1) all 8 Sifrim articles,

as shown above, were not scientifically substantiated; (2) recommendations based on the Delphi method have no scientific status; (3) whatever changes are made at periodically recurring meetings of regular participants, their goal is to advertise pH monitoring. Although they claim that reflux can lead to dilated intracellular spaces, however, believe that this does not always happen. At the same time, Naik and Vaezi concluded that, "The future of reflux diagnosis may very well be without the need for currently employed technologies and could be as simple as assessing changes in epithelia integrity as a surrogate marker for GERD" [16]. This method for diagnosing GER, despite the use of electron microscopy, they called reliable and simple.

Recently, the scientifically proven hypothesis that intestinalization of metaplastic cardiac mucosa is an early sign of GER [12,13] has been completely ignored. It is obvious that the earlier the GER is diagnosed, the more likely the patient will recover, or the progression of the disease will be stopped. Treating heartburn as a pain rather than as a symptom of GER leads to a late diagnosis and chronicity of the process.

Third, the authors' assertion that FH, unlike GERD, is refractory to acid-suppressant treatment is unfounded. For example, in a systematic review and meta-analysis Simadibrata et al. states that "Up to 40% of gastroesophageal reflux disease (GERD) patients experience inadequate symptom relief with a proton pump inhibitor (PPI), termed PPI-resistant or refractory GERD" [18]. It is also known that PPIs are the most common first-line treatment for heartburn symptoms because some patients report some relief. The authors of the peer-reviewed article, referring to Lyon consensus 2.0, suggest that acid-suppressant therapy may indirectly affect esophageal pain modulation. This implies that PPIs are effective in GERD because they reduce the effect of acid on the esophagus. And when acid reflux is less than 1.5 hours per day, PPIs only reduce pain. This assumption is counterintuitive because reflux is present in both GERD and FH, and there is no other option than to consider them the same disease, i.e., GER.

DeMeester et al. proposed the "DeMeester score" as a boundary between normal and GERD based on the absence or presence of heartburn. Now, the authors, using this method, exclude GERD in most patients with heartburn. I do not see the logic in this situation.

Fourthly, The article does not cite any scientific papers. Meanwhile, the study by Weijenborg et al. showed that: - "Patients with FH did not show acid hypersensitivity as seen in patients with NERD. However, once perceived, intensity of heartburn is similar. Esophageal mucosal integrity is similar between NERD and FH patients and is therefore unlikely

to be the underlying cause of the observed difference in esophageal acid perception" [19]. The authors of the peer-reviewed article refer to the Rome IV criteria, in which the diagnostic criteria for FH require that the patient experience a burning sensation behind the breastbone, discomfort or pain at least twice a week for at least 6 months. These criteria, approved by voting at meetings of specially selected gastroenterologists, together with a false pH criterion for diagnosis of FH and non-pathophysiological treatment, lead to a delay in antireflux therapy, which is the result of the chronicity of the process and leads to the development of esophageal cancer [12]. The statement that the risk of progression to esophageal adenocarcinoma is minimal confirms that we are talking about acid reflux. However, it is impossible to judge the degree of risk, since many years must pass for this to happen, while the diagnosis of FH began to be made only in the early 2000s.

Lyon consensus 2.0. The article by Davis and Gyawali shows how the Delphi process at Lyon consensus 2.0 changed the ideas about FH without the use of scientific research. After 5 years, it turned out that symptom relief as a result of using PPI in more than one-third of patients with normal upper endoscopy and reflux monitoring, likely related to placebo effect and/or incomplete GERD evidence on 24-hour reflux monitoring. It might seem that the authors finally acknowledged the failure of pH monitoring, but further analysis shows that this is an advertisement for more expensive equipment - pH impedance monitoring, the results of which are based on the previous understanding of the possibility of physiological reflux.

The defining features of GERD include an abnormal reflux monitoring study and/or findings on upper endoscopy. However, they have limitations. If these studies do not reveal pathology, but objective evidence of GERD has been previously demonstrated on either endoscopy or ambulatory reflux monitoring, then this pathology is defined as GERD. The second limitation is the statement that "low-grade esophagitis (Los Angeles grade A esophagitis) can be seen in healthy asymptomatic individuals and therefore does not constitute conclusive evidence of GERD" [2,20]. It follows that these methods themselves are not accurate, although everywhere, pH monitoring has been declared the defining diagnostic method for GERD, distinguishing it from supposedly functional diseases, including FH.

The functional disorders of the esophagus, which are supposedly disorders of gut-brain interaction (DGBI) are difficult to differentiate from the GERD without proper investigation for ongoing symptoms. Thus, the authors lead readers to the idea of the need to use pH impedance monitoring. At the same time, other research methods

(histological, radiological) are completely dismissed, without any justification and contrary to common sense. For example, the consensus members voted to accept the statement without explaining what other causes cause the low-grade esophagitis and how it was proven that it is not GERD. Contrary to scientific data, the statement was made public that: - "Recent data indicates that histopathology has a low diagnostic yield and only provides helpful clues to an underlying inflammatory mucosal disorder such as eosinophilic esophagitis when presentation consists of dysphagia or food impaction, or when endoscopic findings of eosinophilic esophagitis are found" [2]. This statement also contradicts common sense, because without histology it is impossible to diagnose eosinophilic esophagitis at a stage when there is no dysphagia [21].

Leading organizers of conferences using the Delphi method, Kia et al acknowledge that "From a technical standpoint, as originally described, DIS measurements have been obtained using transmission electron microscopy with good sensitivity and specificity, albeit expensive, not widely available, and time-consuming" [22]. We see a conflict of interests. The goal of science is to determine the boundary of the histological norm for GERD diagnostics, and not for widespread use in practice. If transmission electron microscopy has good sensitivity and specificity, then it can be considered the gold standard for GERD diagnostics. Since histological changes can be observed with a high pH of the stomach, this proves that pH monitoring does not diagnose weakly acidic GER. It follows that all so-called functional diseases of the esophagus are GER, the treatment of which should be comprehensive, as early as possible, regardless of the clinical symptoms. Then there is no need for pH research in the esophagus and perform histological examinations for each patient.

The aim of the review article by Peter J Kahrilas, John E Pandolfino, and Leila Kia, is visible in the following acknowledgement: -"Seminal work by Farré et al first proposed that BI (baseline impedance) correlated with transepithelial resistance, a known marker of esophageal mucosal integrity, based on in-vivo and in-vitro studies of acid perfusion in animals and humans" [22]. This idea became decisive in the diagnosis of GERD in the Lyon consensus 2.0 recommendations [1]. However, from the article by Farré et al it follows that although patients with GORD have a lower impedance baseline than healthy volunteers at the distal esophagus, acidic solutions in all patients induced DIS. Secondly, these changes were observed after a single acid administration, which contradicts the idea of the possibility of physiological reflux [16]. In a later study, these authors showed that approximately 30% of healthy volunteers show dilated intercellular spaces in the esophageal epithelium probably due to increased presence of physiological acid reflux [17].

Finally, the question arises why the authors admit that mucosal damage from reflux leads to dilated intracellular spaces, but do not recommend this method for diagnosing GERD. Involuntarily, the assumption arises that histological examination is a more accurate competitor to pH monitoring, which they devote their articles to advertising. Moreover, if dilated intracellular spaces indicate GERD, then why does overt esophagitis not confirm GERD? This article reviews 135 articles, most of which are similar reviews or lectures by regular participants in various consensus groups receiving financial support from different companies [21,23,24].

CONCLUSION

The decisions at various conferences on the diagnosis and treatment of gastroesophageal reflux, made using the Delphi method, are not scientific. The analysis of the recommendations revealed that they often contradict known scientific facts, contradict each other and sometimes to common sense. Because of using pH monitoring as a gold standard, which it is not, in about two thirds of patients with GER are unreasonably diagnosed with various supposedly functional diagnoses. As a result, they do not receive timely pathogenic treatment, which leads to chronicity of the process and severe complications. A free discussion is needed to eliminate contradictions in modern gastroenterology.

ACKNOWLEDGEMENTS

CONFLICTS OF INTEREST

The author declares that there are no conflicts of interest.

REFERENCES

1. Gabbard S, Vijayvargiya S. (2019). Functional heartburn: An underrecognized cause of PPI-refractory symptoms. Cleve Clin J Med. 86(12):799-806.

2. Davis TA, Gyawali CP. (2024). Refractory Gastroesophageal Reflux Disease: Diagnosis and Management. J Neurogastroenterol Motil. 30(1):17-28.

3. Vakil N, van Zanten SV, Kahrilas P, Dent J, Jones R; Global Consensus Group. (2006). The Montreal definition and classification of gastroesophageal reflux disease: a global evidence-based consensus. Am J Gastroenterol. 101(8):1900-1920.

4. Graham DY. (2007). The Montreal classification: a consensus with single PHARMA sponsor and paid participants. Am J Gastroenterol. 102(2):460-461.

5. DeMeester TR, Johnson LF. (1976). The evaluation of objective measurements of gastroesophageal reflux and their contribution to patient management. Surg Clin North Am. 56(1):39-53.

6. Demeester TR, Johnson LF, Kent AH. (1974). Evaluation of current operations for the prevention of gastroesophageal reflux. Ann Surg. 180(4):511-525.

7. Johnson LF, Demeester TR. (1974). Twenty-four-hour pH monitoring of the distal esophagus. A quantitative measure of gastroesophageal reflux. Am J Gastroenterol. 62(4):325-332.

8. Yoo SS, Lee WH, Ha J, et al. (2007). The prevalence of esophageal disorders in the subjects examined for health screening. Korean J Gastroenterol. 50(5):306-312.

9. Stål P, Lindberg G, Ost A, Iwarzon M, Seensalu R. (1999). Gastroesophageal reflux in healthy subjects. Significance of endoscopic findings, histology, age, and sex. Scand J Gastroenterol. 34(2):121-128.

10. Shieh TY, Chen CC, Chou CK, et al. (2022). Clinical efficacy and safety of peroral endoscopic myotomy for esophageal achalasia: A multicenter study in Taiwan. J Formos Med Assoc. 121(6):1123-1132.

11. Chen JW, Vela MF, Peterson KA, Carlson DA. (2023). AGA Clinical Practice Update on the Diagnosis and Management of Extraesophageal Gastroesophageal Reflux Disease: Expert Review. Clin Gastroenterol Hepatol. 21(6):1414-1421.e3.

12. Labenz J, Chandrasoma PT, Knapp LJ, DeMeester TR. (2018). Proposed approach to the challenging management of progressive gastroesophageal reflux disease. World J Gastrointest Endosc. 10(9):175-183.

13. Chandrasoma P, DeMeester T. (2016). A New Pathologic Assessment of Gastroesophageal Reflux Disease: The Squamo-Oxyntic Gap. Adv Exp Med Biol. 908:41-78.

14. Tseng D, Rizvi AZ, Fennerty MB, et al. (2005). Forty-eight-hour pH monitoring increases sensitivity in detecting abnormal esophageal acid exposure. J Gastrointest Surg. 9(8):1043-1051.

15. Caviglia R, Ribolsi M, Maggiano N, et al. (2005). Dilated intercellular spaces of esophageal epithelium in nonerosive reflux disease patients with physiological esophageal acid exposure. Am J Gastroenterol. 100(3):543-548.

16. Naik RD, Vaezi MF. (2017). Recent advances in diagnostic testing for gastroesophageal reflux disease. Expert Rev Gastroenterol Hepatol. 11(6):531-537.

17. Gorgulu V, Ergun P, Kipcak S, Doganavsargil B, Sifrim D, Bor S. (2023). Revisiting the Role of Esophageal Mucosal Dilated Intercellular Spaces in the Diagnosis and Pathophysiology of Heartburn. J Neurogastroenterol Motil. 29(4):436-445.

18. Simadibrata DM, Lesmana E, Fass R. (2024). A systematic review and meta-analysis of the efficacy of vonoprazan for proton pump inhibitor-resistant gastroesophageal reflux disease. J Gastroenterol Hepatol. 39(5):796-805.

19. Weijenborg PW, Smout AJ, Bredenoord AJ. (2016). Esophageal acid sensitivity and mucosal integrity in patients with functional heartburn. Neurogastroenterol Motil. 28(11):1649-1654.

20. Gyawali CP, Yadlapati R, Fass R, et al. (2024). Updates to the modern diagnosis of GERD: Lyon consensus 2.0. Gut. 73(2):361-371.

21. Levin MD. Eosinophilic Esophagitis: Etiology, Pathogenesis, Diagnosis and Treatment. Hypothesis. DOI: 10.20944/preprints202501.1947.v1

22. Kia L, Pandolfino JE, Kahrilas PJ. (2016). Biomarkers of Reflux Disease. Clin Gastroenterol Hepatol. 14(6):790-797.

23. Farré R, Blondeau K, Clement D, et al. (2011). Evaluation of oesophageal mucosa integrity by the intraluminal impedance technique. Gut. 60(7):885-892.

24. Pauwels A, Broers C, Vanuytsel T, et al. (2018). A reduced esophageal epithelial integrity in a subgroup of healthy individuals increases with proton pump inhibitor therapy. United European Gastroenterol J. 6(4):511-518.