PLAUSIBILITY CRITERIA FOR PUTATIVE PATHOPHYSIOLOGICAL MECHANISMS IN FUNCTIONAL GASTROINTESTINAL DISORDERS: A CONSENSUS OF EXPERTS

Jan Tack¹, Maura Corsetti², Michael Camilleri³, Eamonn MM Quigley⁴, Magnus Simren⁵, Hidekazu Suzuki⁶, Nicholas J. Talley⁷, Hans Tornblom⁵, Lukas Van Oudenhove¹

1. TARGID, University of Leuven, Herestraat 49, 3000 Leuven, Belgium
2. National Institute for Health Research, Nottingham Digestive Diseases Biomedical Research Centre, Nottingham University Hospitals NHS Trust, University of Nottingham, Nottingham, United Kingdom
3. CENTER, Mayo Clinic, Rochester, MN 55905, USA
4. Lynda K and David M Underwood Center for Digestive Disorders, Division of Gastroenterology and Hepatology, Houston Methodist Hospital and Weill Cornell Medical College, Houston, Texas, USA
5. Department of Internal Medicine & Clinical Nutrition, Institute of Medicine, Sahlgrenska Academy, University of Gothenburg, Sweden
6. Medical Education Center, Keio University School of Medicine, 35 Shinnanomachi, Shinjuku-ku, Tokyo, 160-8582, Japan
7. Faculty of Health and Medicine, University of Newcastle, Australia

Corresponding author: Jan Tack, MD, PhD
TARGID, University of Leuven
Herestraat 49, box 701
3000 Leuven, Belgium
jan.tack@med.kuleuven.be

Author contributions:
JT, MS, HT, LVO, MC, MC, EQ, HS, NT: concept, content, manuscript writing, manuscript revision

Conflicts of interest for this manuscript: none
Support: this consensus process was provided by a Methusalem grant from Leuven University to JT.
Abstract

Background and aims: The functional gastrointestinal disorders (FGIDs), are extremely common conditions associated with a considerable personal, social and health economic burden. Managing FGIDs in clinical practice is challenging because of the uncertainty of symptom-based diagnosis, the high frequency of overlap between these conditions and the limited efficacy of available therapies. It has often been argued that successful drug development and management of FGIDs requires knowledge of the underlying pathophysiology. Numerous and highly variable candidate pathophysiological mechanisms have been implicated in the generation of FGID symptoms, but there is no current consensus on how to best define the relevance of these disturbances. Methods: A group of international experts on FGIDs developed plausibility criteria that should be fulfilled by relevant pathophysiological mechanisms in FGIDs. Results: Five criteria are proposed: 1) presence of the abnormality in a subset of patients; 2) temporal association between proposed mechanism and symptom(s); 3) correlation between the level of impairment of the mechanism and symptom(s); 4) induction of the symptom(s) by provoking the pathophysiological abnormality in healthy subjects and 5) treatment response by a therapy specifically correcting the underlying disorder, or congruent natural history of symptoms and dysfunction in the absence of specific therapy. Based on strength of evidence for these 5 criteria, a plausibility score is proposed. Conclusion: Evaluation of the strength of evidence for candidate pathophysiological abnormalities fulfilling these 5 plausibility criteria will help to identify the most relevant mechanisms to target for novel diagnostic approaches and for the development of new therapies.

Keywords: functional gastrointestinal disorders, pathophysiological mechanisms, gastrointestinal motility, gastroesophageal reflux, visceral hypersensitivity, plausibility criteria
Summary box
What is already known about this subject:

- Several pathophysiological mechanisms have been implicated in the generation of functional gastrointestinal disorders
- Variable arguments to support the relevance of individual candidate pathophysiological mechanisms
- No standard set of requirements has been defined that would establish the relevance of an individual candidate pathophysiological mechanism

What are the new findings:

- A group of international experts developed plausibility criteria for mechanisms in functional gastrointestinal disorders
- These are based on aspects such as presence, temporal association, correlation between level of impairment and symptom severity, induction in healthy subjects and treatment response or congruent natural history
- The plausibility criteria were applied to 4 specific mechanisms in 3 different functional disorders

How might it impact on clinical practice in the foreseeable future?

- Evaluation of plausibility criteria allows 1) to summarize current knowledge, 2) to identify gaps requiring targeted future research and 3) may help to advance the field of functional gastrointestinal disorders
1. Introduction.

Functional gastrointestinal disorders (FGIDs) are characterized by the presence of a variety of chronic, typically episodic symptoms attributed to the gastrointestinal tract in the absence of an underlying histological, biochemical, or physiological mechanism that consistently explains the symptoms (1). FGIDs may be characterized by symptoms that are attributable to the pharynx, esophagus, stomach, biliary tract, small intestine, colon or anorectum, and their pathogenesis has remained obscure (1). However, over the last two decades, it has become clear that a significant proportion of FGIDs are associated with at least one pathophysiological abnormality, each being observed in 25 to 40% of the patients. Thus, a number of putative pathophysiological mechanisms have been proposed, including disordered motility, visceral hypersensitivity, low-grade inflammation, altered microbiota, immune activation, adverse reactions to foods and central nervous system dysfunction (which may or may not be related to psychological dysfunction) (2-6).

Despite the fact that these disturbances have been reported in patients with FGIDs, their relevance to symptom generation remains the subject of debate, in part because of the absence of a clearly established causal or even temporal relationship between symptoms and observed abnormal function, as well as the inadequacy of medications to specifically target the putative underlying mechanisms.

Several cross-sectional studies attempting to correlate symptoms with pathophysiological mechanisms in FGIDs have been criticized because they failed to explain a given symptom in all patients, or because of an inability to rule out other contributing mechanisms (7-10). This controversy reflects a number of specific problems associated with FGIDs. In functional syndromes, the assessment of the nature and the severity of symptoms is currently dependent on patient self-reports. In addition, it is often assumed that FGIDs consist of subgroups with heterogeneous symptoms and different underlying pathophysiology (11-13).

2. The case of heartburn and gastro-esophageal reflux of acid.
To illustrate the concept of developing criteria to establish relationships between pathophysiological mechanism(s) and gastrointestinal disorders, let us consider the role of acid gastro-esophageal reflux in the pathophysiology of heartburn. The broad acceptance of this mechanism of symptom generation as being causal is based on several observations: first, the finding of increased esophageal acid exposure on pH monitoring in patients with heartburn (“presence”); second, the temporal association between reflux events and heartburn occurrence (“temporal association”); third, the worsening of heartburn scores with increasing severity of esophageal acid exposure (“correlation”); fourth, the induction of heartburn by esophageal acid perfusion (“induction”) and fifth, the response of symptoms to acid-suppressive therapy (“treatment response”) (14). The pathophysiological relevance of acid reflux for heartburn symptom generation is well established in spite of the fact that not all heartburn can be attributed to (acid) reflux, with these patients being referred to as having functional heartburn (14,15). Conversely, not all subjects with increased esophageal acid exposure experience symptoms of heartburn (16), indicating that 100% correlation is not considered necessary.

To date, no criteria have been proposed to establish plausibility for a causal relationship between a pathophysiological mechanism and one or more symptoms in FGIDs. In infectious disease, for instance, Koch’s postulates are well-accepted, representing rigorous criteria that link infectious diseases to specific microbial infectious agents (17). Similarly, the Bradford Hill criteria are a group of criteria that provide evidence of a causal relationship between a putative risk factor and a possible consequence in epidemiology research (18). Research on pathophysiological mechanisms of FGIDs may benefit and gain focus if a similar set of standards could be proposed.

The aim of this paper is to develop a set of standards to be used as criteria for qualifying putative pathophysiological mechanisms as relevant to the symptoms and management of patients with FGIDs.

3. Consensus development process.
The decision to initiate a consensus effort to establish plausibility of mechanisms underlying FGID symptom generation was taken by a small core group (JT, MS, HT, LVO) that prepared and initiated the process after a meeting in Gothenburg in August 2014. As the issue of plausibility criteria for pathophysiological mechanisms has no existing precedent efforts, and is not readily amenable to traditional scientific investigation, an approach based on expert-based consensus was chosen (19). The core group identified a group of eligible contributors, with a well-established track record in studying FGID pathophysiology, who were selected and invited to participate in the process. The selection process took into account 1) diversity in pathophysiology research expertise (gastrointestinal motility, visceral hypersensitivity, microbiota, psychosocial co-morbidities, central nervous system processes), 2) focus on specific parts of the GI tract (esophagus, stomach, bowel), 3) affiliations with different societies (UEG, AGA, WGO, European, American and Asian-Pacific Motility societies) and 4) region. All invited participants agreed to participate and were supportive of the scope and general aims of the project, leading to the consensus group comprising the current 9 members.

The core group drafted a number of plausibility criteria which were described in a working document. Face-to-face interactions of (subgroups of) participants occurred at international conferences in 2014 and 2015. These interactions served to adapt, expand or decrease the list of plausibility criteria based on full consensus, and to generate a list of mechanisms and conditions to be considered for illustrative analysis. These included diverse organ systems and candidate pathophysiological mechanisms. Proposed topics included esophageal hypersensitivity in functional heartburn and non-cardiac chest pain, reflux in non-cardiac chest pain, gastric emptying in functional dyspepsia (FD), impaired gastric accommodation in FD, H. pylori infection in FD, gastric hypersensitivity in FD, increased mucosal permeability in the irritable bowel syndrome (IBS), transit disorders in IBS, rectal hypersensitivity in IBS, altered microbiota in IBS and stress and anxiety in FD or IBS, amongst others.

The consensus process generated a final list of 6 different plausibility criteria, and selected 5 combinations of pathophysiological mechanisms and FGIDs for potential illustrative analysis. The members of the consensus group drafted different text segments which were merged into a first draft of the manuscript which was circulated.
amongst all participants in the Summer of 2015. All members commented on content and wording, and areas of disagreement were identified, discussed and, if needed, adapted. This process resulted in an optimization of the plausibility criteria wording, a restructuring of the criteria into 5 categories (one with two components) and an ongoing discussion on the number of illustrative analyses to be included. An updated version of the criteria was generated and circulated for evaluation early in 2016, and all panel members agreed in full with the current set of criteria. Based on the discussion, strength of evidence grading was also added to the document (20), and the illustrative analyses were chosen, taking into account the desire to include variations in type of pathophysiological mechanism and different organs of the gastrointestinal tract, while avoiding topics that were recently evaluated in detail in the literature (21). Revisions throughout 2016 led to finalization of the current document.


Based on the case of heartburn and acid reflux, a number of “criteria” can be put forward that add weight to accepting a causal relationship between a pathophysiological mechanism and a (functional) gastrointestinal symptom. These include presence of the mechanism in at least a subset (if not all) of the subjects with the symptom (criterion 1: “presence”), a temporal association between the presence of the mechanism and symptom occurrence (criterion 2: “temporal association”); a correlation between the level of impairment of the mechanism and the severity of the symptom, which is usually cross-sectional (criterion 3: “correlation”); induction of the symptom by inducing the pathophysiological abnormality in asymptomatic subjects (criterion 4: “induction”); and, finally, a favorable response to a treatment aimed at specifically correcting the underlying pathophysiological mechanism (“criterion 5: treatment response”) (Table 1). In the absence of a specific therapeutic intervention, observations of congruent evolution over time between the severity of the symptom and the severity of the pathophysiological mechanism (criterion 5b: “congruent natural history”) is a potential alternative approach. The strength of evidence can be evaluated using the GRADE system (Table 2) (20). Similar to what was pointed out for subjects with pathological acid exposure without heartburn symptoms, the
presence of a putative pathophysiological abnormality in some subjects in the absence of symptoms does not necessarily invalidate its relevance.

In the following paragraphs, we will evaluate the extent to which four frequently proposed mechanisms in three common FGIDs fulfill these criteria. We chose to evaluate conditions in different anatomical regions of the gastrointestinal tract, and we selected a disorder of chemosensitivity (acid reflux in non-cardiac chest pain), two disorders of motility control (delayed gastric emptying and impaired gastric accommodation in functional dyspepsia (FD)) and a disorder of mechanosensitivity (colorectal hypersensitivity in IBS) as primary examples. These examples are not chosen because the mechanisms are unequivocally correlated with symptom(s) on all 5 criteria. Rather, they were chosen because the published evidence illustrates some positive criteria, and also because they fail to fulfill some of the plausibility criteria, thereby identifying areas requiring future research for the validation of the putative mechanism.

5. Relationship between symptoms and acid reflux in non-cardiac chest pain.

Presence: Esophageal pH monitoring demonstrates the presence of pathological esophageal acid exposure in up to 62% of the patients with non-cardiac chest pain (22). However, even in the absence of pathological reflux, reflux can still be the cause of a patient’s symptoms of chest pain in patients with an acid-hypersensitive esophagus (23).

Temporal association: In non-cardiac chest pain, esophageal pH monitoring with symptom registration shows a significant temporal association between acid reflux and symptom occurrence in, on average, 31% of all patients with this symptom (23).

Correlation: In population-based studies, the prevalence of chest pain of presumed non-cardiac origin increased with increasing frequency of typical GERD symptoms (which are in turn assumed to be caused by acid reflux) (24,25). Whether the severity of chest pain increases with increasing acid exposure in NCCP patients has not been adequately assessed.

Induction: Esophageal acid perfusion mimics symptoms in 10 to 38% of unselected patients with non-cardiac chest pain (22). However, there are no reports of induction of reflux-related chest pain in healthy volunteers.
Therapeutic response: Case series and controlled trials demonstrate that proton pump inhibitor therapy may be effective in up to 80% of patients selected for reflux-related non-cardiac chest pain and its empirical use for both diagnostic and therapeutic purposes has been advocated for this reason (22,23).

Congruent natural history: No longitudinal follow-up studies evaluating relationships between acid reflux and symptoms of non-cardiac chest pain are available outside of the setting of interventional therapeutic trials. The fulfillment of criteria for association of acid reflux with non-cardiac chest pain is summarized in Table 3.

6. Relationship between symptoms and impaired gastric accommodation in functional dyspepsia.

Presence: A number of studies, using different methods, reported that meal-induced gastric accommodation is impaired in approximately 40% of FD patients (2).

Temporal association: FD is increasingly considered to be a meal-related condition (26) and this is especially true for the subgroup referred to as the postprandial distress syndrome, according to the Rome III definition (27). In a large study in tertiary care, symptoms were triggered by ingestion of a meal in the vast majority of FD patients, reached a maximum early after ingestion and gradually decreased over the 4-hour measurement period (26). The time course of these symptoms parallels the arrival and presence of food in the stomach, and the process of gastric accommodation, which allows the proximal stomach to receive and store food. Nevertheless, other data suggest that some FD symptoms are also generated from the small intestine (28).

Correlation: Studies evaluating overall FD symptom severity failed to find a significant impact of the presence of impaired accommodation (2,29,30). On the other hand, a number of studies reported associations between impaired gastric accommodation and the presence and severity of early satiation (31,32) or with the presence of unexplained weight loss (32,33), but others failed to find such association (34). Thus, while there is some evidence for an association with the symptoms of early satiation and weight loss, the severity of other symptoms of FD does not seem to be correlated to gastric accommodation.
**Induction:** Impaired accommodation was induced in healthy volunteers through pharmacological interventions (nitric oxide synthase inhibition, motilin receptor agonists, opioid receptor antagonism, GLP-1 analogue) or by psychological manipulation (anxiety induction), and this was associated with early satiation and decreased nutrient volume tolerance (35-39). In contrast, administration of both the cannabinoid (CB)1 receptor antagonist rimonabant and of the gut peptide ghrelin inhibited gastric accommodation in healthy controls, while no significant difference in meal-induced satiation occurred but some other symptom ratings tended to increase (40,41). Taken together, the majority of studies showed that induction of impaired accommodation also induced early satiation in healthy volunteers.

**Therapeutic response:** A number of drugs with the potential to enhance gastric accommodation have been evaluated in FD patients, but most of them have several additional actions and the link between symptom response and change in accommodation is often not assessed. A pilot study with the 5-HT1A agonist buspirone showed a relationship between enhancement of gastric accommodation and improvement of symptoms in a small group of FD patients (42). A larger multicenter study in Japan with tandospirone, also a 5-HT1A agonist, confirmed efficacy in FD that could not be attributed to the anxiolytic and antidepressant properties of the drug (43).

Acotiamide is a novel, first-in-class drug which is an antagonist of the inhibitory muscarinic type 1 and type 2 (M1/M2) auto-receptors on cholinergic nerve endings, and is also a cholinesterase inhibitor (44). Through these actions, acotiamide enhances the availability of synaptically released acetylcholine, thereby enhancing reflex-controlled motility. In a placebo-controlled pilot study, acotiamide was superior in providing symptom relief in FD, and this was accompanied by enhancement of gastric accommodation (45). The enhancing effect of acotiamide on gastric accommodation was confirmed using ultrasound in a small study in Japan (46). Based on phase 2 and 3 studies, the drug was approved in Japan for the treatment of FD (47-49). However, the drug also enhances gastric emptying rate and inhibits stress-related effects on gastric function and nutrient intake (46,50), and thus impaired accommodation is not necessarily the only mechanisms whereby acotiamide may exert its therapeutic effect.

**Congruent natural history:** To our knowledge, no longitudinal follow-up studies assessing evolution of both symptoms and gastric accommodation in FD are available.
The criteria for association of impaired gastric accommodation with FD symptoms are summarized in Table 3. While some criteria are fulfilled, it is clear that there is a need for additional studies on the strength of the association between FD symptoms and impaired accommodation, and the evolution of this association, if any, over time.

7. Relationship between symptoms and delayed emptying in functional dyspepsia.

Presence: Impaired gastric emptying has been proposed as one of the underlying pathophysiological mechanisms in FD. Several studies reported that gastric emptying of a radiolabeled meal was delayed in 25 to 50% of FD patients (2). Notwithstanding the occurrence of delayed emptying in FD, it cannot completely account for all symptoms since a major proportion of patients have a normal emptying rate.

Temporal association: Only a few studies have addressed the occurrence of FD symptoms in relation to gastric emptying. In a large study in tertiary care FD patients, symptoms were triggered by ingestion of a meal in the vast majority of patients, reached a maximum early after ingestion and gradually decreased over the 4-hour measurement period (26). While the time course of these symptoms parallels the presence and emptying of food from the stomach, a recent study also provided evidence that some FD symptoms are also likely to be generated from the small intestine given their time course relative to meal ingestion (51).

Severity correlation: Several studies have evaluated the relationship between delayed gastric emptying and FD symptom pattern and severity. While smaller sized studies failed to find an association, a number of studies which included more than 300 patients found associations between delayed emptying and the symptoms postprandial fullness, nausea or vomiting (2,52-55). These large-scale studies reported higher prevalence rates for specific symptoms in FD patients with delayed emptying compared to those with normal emptying. The correlation between the severity of these symptoms and the degree of delayed emptying was not reported in detail. Hence, the strength of the correlation between delayed emptying and dyspepsia symptom severity is understudied.

Induction: Several studies have used nutritional, physical or pharmacological interventions to delay gastric emptying rates in healthy subjects, but none assessed the occurrence of dyspepsia-like symptoms in the participants (56-65). In a study
using sumatriptan as a pharmacological agent to delay gastric emptying in healthy subjects, no increase in meal-related symptoms was induced (66).

**Therapeutic response:** Gastroprokinetic drugs (such as cisapride, domperidone, erythromycin and tegaserod), have been extensively evaluated in clinical trials in FD patients. In meta-analyses of the published literature, prokinetics were superior to placebo in providing relief of FD symptoms, but there were indications of publication bias with omissions of negative trials (67). Furthermore, the symptomatic benefit of prokinetics was poorly correlated to their effect on gastric emptying rate (68), suggesting that effects of these drugs on gastric accommodation, sensitivity to gastric distention and on gastro-esophageal reflux may all contribute to the symptom impact in FD (69,70). In keeping with this hypothesis, large clinical trials with the potent prokinetic macrolide ABT-229 were unequivocally negative with regard to symptom improvement (71).

**Parallel evolution:** To our knowledge, no longitudinal follow-up studies assessing the evolution of both symptoms and gastric emptying in FD are available. The criteria for association of delayed gastric emptying with FD symptoms are summarized in Table 2. While some criteria are fulfilled, it is clear that there is room for additional studies on the strength of the association between symptom severity and delayed emptying, and its evolution over time.

8. Relationship between symptoms and colorectal hypersensitivity in the irritable bowel syndrome.

**Presence:** Colorectal hypersensitivity is considered to be one the most important pathophysiological mechanisms in IBS Several studies have confirmed that IBS patients show increased sensitivity to colorectal balloon distention, while colorectal compliance is not altered (72-83). Colorectal hypersensitivity is present in a large subset of IBS patients, but it is not present in all, so it cannot account for all IBS symptom generation.

**Temporal association:** The temporal relationship between IBS symptom generation and colorectal hypersensitivity is only partly established. In fact, colorectal hypersensitivity is often considered a marker for more generalized intestinal hypersensitivity, and it is thought that the entire gastrointestinal tract may contribute to
symptom generation (84,85). The strongest indication of a temporal association between colorectal sensitivity and symptoms is derived from studies of meal aggravation of IBS symptoms: the increase of IBS symptoms after a meal paralleled increased rectal sensitivity (86-88), although healthy controls were also sensitized after meal ingestion (89).

**Correlation:** A number of large-scale studies reported associations between colorectal hypersensitivity and the presence and severity of abdominal pain and other symptoms in IBS (73,78,80,82,83), but the strength of the correlation was weak to moderate at best, and there are also studies that revealed no association between colorectal hypersensitivity and overall IBS symptom severity (75,89). In fact, psychologically driven severity reporting has been implicated as a confounder of assessment of sensitivity to colorectal distention in IBS (90,91).

**Induction:** Only one group reported on induction of rectal hypersensitivity in healthy volunteers; in that instance by the administration of glycerol (92,93). Attempts to sensitize the rectum using tryptophan depletion, antagonism of endogenous opioid pathways or anxiety induction, were not successful (94-97). In post-hoc analysis, acute tryptophan depletion increased perception ratings to distention in the lower pressure ranges, but not in the more intense and pain-generating ranges (95). Nutrients induced increased rectal sensitivity and this was associated with more severe symptoms in IBS (86-88), and to some extent also in healthy individuals (89).

**Therapeutic response:** A number of studies, generally of small size, have evaluated changes in colorectal sensitivity induced by drugs known to provide symptomatic benefit in IBS patients. The clinical efficacy of alosetron in IBS-D was not paralleled by a change in sensitivity thresholds to isobaric visceral distention in a separate mechanistic study (98,99). Similarly, the clinical efficacy of tegaserod, as well as lubiprostone, in IBS-C was not associated with increases in thresholds to elicit discomfort or pain during rectal sensitivity testing in separate smaller studies (98,100-102). In studies of psychological therapies in IBS (cognitive-behavioral therapy or brief psychotherapy), symptom improvement was paralleled by improvements in visceral sensitivity and the same has also been found for treatment with antidepressants in some, but not all, studies (103-108). Moreover, use of intrarectal lidocaine has, in small studies, reversed rectal hypersensitivity and improved abdominal pain in patients with IBS (109,110). Ketotifien improved symptoms of abdominal pain in a controlled study in IBS, and also improved visceral hypersensitivity, but there was no
correlation between changes in sensory thresholds and changes in abdominal pain scores after treatment (111).

**Congruent natural history:** Conflicting results exist regarding the evolution of colorectal sensitivity over time in IBS patients and its association with symptom evolution. In one study a small group of IBS patients (n=15) underwent rectal balloon distensions and GI symptom assessment on two occasions, three months apart, and a significant correlation between changes in IBS symptoms and rectal sensitivity was demonstrated (83). Another study assessed evolution of rectal sensitivity over 12 weeks in 33 IBS patients and found that rectal sensitivity was either unchanged or increased in patients with stable GI symptoms (78). In contrast, a longitudinal follow-up study applied rectal sensitivity evaluations to IBS patients and healthy controls at 4-month intervals for one year. In this study, perceptual ratings in response to rectal distention returned to control levels over 12 months, whereas IBS symptom severity did not significantly change (112). This suggests that part of the increased sensitivity to rectal distention in IBS may be related to stress and anxiety related to the index procedure, and that repetition of the stimulus results in habituation to visceral distention tests.

The criteria for association of colorectal hypersensitivity with IBS symptoms are summarized in Table 3. While some criteria are fulfilled, it is clear that there is a need for additional studies on the strength of the association between IBS symptoms and colorectal hypersensitivity, and its relevance as a therapeutic target.
8. Plausibility score

Based on the strength of evidence as evaluated using the GRADE system (20), a numerical plausibility score can be constructed, by allocating a score ranging from 3 to 0, corresponding to GRADE ratings of A to D respectively, for the first 4 criteria (“Presence”, “Temporal association”, “Correlation” and “Induction”). For the 5th criterion items (“Therapeutic response” or “Congruent natural history”), the highest of both is used to generate a 5th number; adding up the five numbers generates a plausibility score between 0 and 15. Applying these scores to the pathophysiological mechanisms listed in Table 3 generates plausibility scores of 11 of acid reflux in NCCP, of 9 for impaired accommodation in FD, of 7 for delayed emptying in FD and of 6 for rectal hypersensitivity in IBS (Table 3).

9. Discussion

The purpose of this manuscript was to propose a number of criteria that should be fulfilled before a pathophysiological mechanism can be considered to be involved in symptom generation in a FGID. We used the well-accepted association between heartburn and acid reflux as the benchmark and, based on the state of knowledge on this association and interactions between the authors of this manuscript, we developed a set of 5 plausibility criteria. These include obvious aspects like criterion 1: “presence”, criterion 2: “temporal association” and criterion 3: “correlation”, but also additional requirements such as criterion 4: “induction” and “criterion 5: treatment response”.

It is probably true that most researchers already use a type of subjective plausibility evaluation when formulating or pursuing hypotheses, and hence many may believe that such a set of criteria already exists. However, no publication proposing plausibility criteria exists in the literature, and the interactions amongst the panel clarified that members had a number of plausibility criteria driving their own pathophysiological research. However, none of the participants alone was considering all criteria proposed in this manuscript, again confirming the value of the consensus approach.
There are obvious limitations to the feasibility of studying each of these criteria in the context of a given potential pathophysiological mechanism. For instance, induction of the abnormality in healthy volunteers may not always be possible, as this depends on the nature of the abnormality and the availability of suitable interventions to mimic it temporarily in an innocuous way. Treatment options considered for criterion 5 aim at providing symptomatic relief by correcting a pathophysiological mechanism. There are evident limitations to criterion 5, as suitable and selective treatment options are lacking. In this case, criterion 5b: “congruent natural history” may serve as a valid, although difficult to accomplish, alternative approach.

The field of FGIDs is often approached with skepticism driven by the assumption that single pathophysiological mechanisms are unlikely to be relevant and that the underlying disease mechanisms are much more complex and defy a reductionist approach. Even if this would be the case, there is a clear need for a structured approach, to take better stock of what is currently known and plausible, to identify areas of uncertainty or lack of data, and to address these in future research agendas based on the plausibility criteria evaluation. The plausibility criteria as described here allow such a systematic approach. Transforming the GRADE strength of evidence into numbers generates a tangible numerical “plausibility score”. This allows comparative evaluations of strength of plausibility for different mechanisms in the same condition, or to compare plausibility of mechanisms across diagnostic entities. In the examples outlined above, the plausibility of disorders explaining a functional disorder is explored. However, the process may also be applied to individual symptoms within a functional disorder (e.g. impaired accommodation explaining early satiation in functional dyspepsia).

Furthermore, even if FGIDs do not depend on single pathophysiological mechanisms, but rather occur in the presence of two or more disease mechanisms which may be closely related and interdependent (e.g. disordered motor and sensory function of stomach or rectum) or unrelated (e.g. a combination of a motor disorder with anxiety), the plausibility evaluation process can be applied to combinations of mechanisms as well. Another limitation is the well-known overlap of FGIDs, which may or may not be caused by common pathogenetic and pathophysiological mechanisms. The current
plausibility criteria are not only applicable to FGIDs as discrete entities but may potentially also be used to evaluate pathophysiological mechanisms underlying overlapping conditions. However, these latter approaches are considerably more complex.

The four mechanisms that were evaluated in the current manuscript cover only a small fraction of all putative pathophysiological mechanisms implicated in the pathogenesis of many different FGIDs. Outside the scope of the present paper, but of major value would be the evaluation of plausibility as mechanisms for many diverse putative pathophysiological factors, such as genetic factors, environmental influences, trauma or abuse, increased mucosal permeability, gut microbiota, etc. Evidently, the number and type mechanisms that can be evaluated depend on the state of the science and hence, over time, seemingly well-accepted mechanisms can be dropped based on new data, completely new mechanisms can emerge based on scientific progress, and old mechanisms can regain interest based on new data or analyses. Scenarios and implications may differ according to the state of science, and hence may change over time. The numerical “plausibility strength” score should also be interpreted taking into account the state of knowledge and evidence. For instance, a low score based on a lack of studies implies that more research is needed, rather than that the mechanism should not be further pursued. Conversely, a situation where a lot of research with mostly negative findings exists argues against the plausibility of a mechanism.

Pending future progress in the area, we propose a framework to evaluate the state of existing knowledge on the role of putative pathophysiological mechanisms in the generation of one or more symptoms in FGIDs. These plausibility criteria, which are easily translated into a numerical score, allow us to summarize current knowledge and to identify gaps requiring targeted future research, which may help to advance the field of FGIDs.
**TABLES**

**Table 1.** Listing of putative criteria for implicating pathophysiological mechanisms in symptom generation of functional gastrointestinal disorders.

<table>
<thead>
<tr>
<th>Number</th>
<th>Criterion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>“Presence”</td>
<td>The pathophysiological disturbance is present in at least a subset of patients with the functional gastrointestinal symptom, and the prevalence is higher than in appropriate controls.</td>
</tr>
<tr>
<td>2</td>
<td>“Temporal association”</td>
<td>A close temporal association exists between the pathophysiological disturbance and the symptom occurrence.</td>
</tr>
<tr>
<td>3</td>
<td>“Correlation”</td>
<td>A significant correlation exists between the presence/severity of the symptom and the presence and the severity of the disturbed function.</td>
</tr>
<tr>
<td>4</td>
<td>“Induction”</td>
<td>Inducing or mimicking the dysfunction generates the symptom in asymptomatic subjects.</td>
</tr>
<tr>
<td>5</td>
<td>A. “Therapeutic response”</td>
<td>Treatment aimed at correcting the underlying disorder improves the symptom. Or:</td>
</tr>
<tr>
<td></td>
<td>B. “Congruent natural history”</td>
<td>Changes in symptom severity parallel changes in the severity of the disturbance.</td>
</tr>
</tbody>
</table>
Table 2. Grade system (ref. 18)

<table>
<thead>
<tr>
<th>Code</th>
<th>Quality of Evidence</th>
<th>Definition</th>
</tr>
</thead>
</table>
| A    | High                | Further research is very unlikely to change our confidence in the estimate of effect.  
• Several high-quality studies with consistent results  
• In special cases: one large, high-quality multi-center trial |
| B    | Moderate            | Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
• One high-quality study  
• Several studies with some limitations |
| C    | Low                 | Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
• One or more studies with severe limitations |
| D    | Very Low            | Any estimate of effect is very uncertain.  
• Expert opinion  
• No direct research evidence  
• One or more studies with very severe limitations |
Table 3. State of fulfillment of putative criteria for implicating pathophysiological mechanisms in symptom generation in 3 prototype functional gastrointestinal disorders, based on the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system (see Table 2).

<table>
<thead>
<tr>
<th>Number</th>
<th>Criterion</th>
<th>NCCP: acid reflux</th>
<th>FD: impaired accommodation</th>
<th>FD: delayed gastric emptying</th>
<th>IBS: rectal hypersensitivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Presence</td>
<td>A</td>
<td>A</td>
<td>A</td>
<td>A</td>
</tr>
<tr>
<td>2</td>
<td>Temporal association</td>
<td>A</td>
<td>B</td>
<td>B</td>
<td>A</td>
</tr>
<tr>
<td>3</td>
<td>Correlation</td>
<td>A</td>
<td>C*</td>
<td>C**</td>
<td>C***</td>
</tr>
<tr>
<td>4</td>
<td>Induction</td>
<td>D</td>
<td>B</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>5</td>
<td>A. Therapeutic response</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td></td>
<td>B. Congruent natural history</td>
<td>D</td>
<td>D</td>
<td>D</td>
<td>D</td>
</tr>
</tbody>
</table>

Plausibility score****: 11 9

* Only supportive data for early satiation and weight loss.
** Only support for postprandial fullness, nausea and vomiting in selected studies
*** A majority of studies found correlations, but they are often in the weak to moderate range and relate to a single symptom or a subset of symptoms.
**** This score is generated by allocating numbers ranging from 3 to 0 for GRADE scores A to D respectively for each of the 4 items and for the highest sub-item 5, and adding these up to generate a total between 0 and 15.
REFERENCES


