Gastroenterology 2016;150:1469–1480

Design of Treatment Trials for Functional Gastrointestinal Disorders

E. Jan Irvine,[1,2,*] **Jan Tack**,[3,*] Michael D. Crowell,[4] Kok Ann Gwee,[5] Meiyun Ke,[6] Max J. Schmulson,[7] William E. Whitehead,[8] and Brennan Spiegel[9]

[1]Department of Medicine, University of Toronto, Toronto, Ontario, Canada; [2]Li Ka Shing Knowledge Institute and Department of Medicine, St Michael's Hospital, Toronto, Canada; [3]Departments of Clinical and Experimental Medicine and Gastroenterology, Translational Research Center for Gastrointestinal Disorders, University Hospital KU Leuven, Leuven, Belgium; [4]Division of Gastroenterology and Hepatology, Mayo Clinic, Scottsdale, Arizona; [5]Yong Loo Lin School of Medicine, National University of Singapore, Singapore; [6]Peking Union Medical College Hospital, Center of FGID and MGID, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China; [7]Facultad de Medicina, Universidad Nacional Autónoma de México, Laboratorio de Hígado, Páncreas y Motilidad, Unidad de Investigación en Medicina Experimental, Hospital General de México, Mexico City, Mexico; [8]University of North Carolina at Chapel Hill, Chapel Hill, North Carolina; and [9]Cedars-Sinai Health System, Cedars-Sinai Center for Outcomes Research and Education, Los Angeles, California

This article summarizes recent progress and regulatory guidance on design of trials to assess the efficacy of new therapies for functional gastrointestinal disorders (FGIDs). The double-masked, placebo-controlled, parallel-group design remains the accepted standard for evaluating treatment efficacy. A control group is essential, and a detailed description of the randomization process and concealed allocation method must be included in the study report. The control will most often be placebo, but for therapeutic procedures and for behavioral treatment trials, respectively, a sham procedure and control intervention with similar expectation of benefit, but lacking the treatment principle, are recommended. Investigators should be aware of, and attempt to minimize, expectancy effects (placebo, nocebo, precebo). The primary analysis should be based on the proportion of patients in each treatment arm who satisfy a treatment responder definition or a prespecified clinically meaningful change in a patient-reported outcome measure. Data analysis should use the intention-to-treat principle. Reporting of results should follow the Consolidated Standards for Reporting Trials guidelines and include secondary outcome measures to support or explain the primary outcome and an analysis of harms data. Trials should be registered in a public location before initiation and results should be published regardless of outcome.

Keywords: Functional Gastrointestinal Disorders; Controlled Trial; Patient-Reported Outcome Measure; Intention to Treat.

Clinical trial design for functional gastrointestinal disorders (FGIDs) is hampered by several factors, including symptom variability between subjects or groups and within subjects over time and the lack of specific biomarkers. The Rome diagnostic criteria and design recommendations are now routinely applied in clinical treatment trials. Since the publication of the Rome III guidance, there have been substantial advances in several aspects of clinical trial design. The expectations for patient-reported outcome (PRO) measurement have undergone major changes with the dissemination of regulatory guidelines for PROs from the US Food and Drug Administration (FDA) and the European Medicines Agency (EMA).[1–3] Accumulating data also provide new insights for measuring common FGID symptoms, such as abdominal pain, discomfort, diarrhea, urgency, constipation, and bloating, among others. New information about the placebo, "nocebo," and "precebo" responses also challenges researchers to consider the biases inherent in FGID trials. In addition, advances in pragmatic clinical trial (PCT) design offer new approaches to measuring the effectiveness of FGID therapies in the context of everyday clinical practice. This updated Rome IV chapter now addresses each of these new trends, provides guidance for investigators seeking to develop and conduct FGID clinical trials, and emphasizes evolving concepts about how best to test the risks and benefits among the full range of FGID treatments.

Identifying the Hypotheses and Research Questions

The first task is to establish the hypothesis of the putative effect of the studied treatment, based on its expected mechanism of action, which generates the specific research question(s) for the proposed trial. As multiple factors contribute to the pathogenesis of FGIDs, it is likely that no single therapeutic approach will fully abolish all symptoms.

*Authors share co-first authorship.

Abbreviations used in this paper: BSFS, Bristol Stool Form Scale; CONSORT, Consolidated Standards for Reporting Trials; EMA, European Medicines Agency; FDA, US Food and Drug Administration; FGID, functional gastrointestinal disorder; HRQOL, health-related quality of life; IBS, irritable bowel syndrome; IBS-C, irritable bowel syndrome with constipation; IBS-D, irritable bowel syndrome with diarrhea; IBS-M, irritable bowel syndrome with constipation and diarrhea; MCID, minimally clinical important difference; PCT, pragmatic clinical trials; PRO, patient-reported outcome; RCT, randomized controlled trial.

Most current article

0016-5085/$36.00
http://dx.doi.org/10.1053/j.gastro.2016.02.010

Table 1. Goals of a Treatment Trial

To ascertain the ability of the intervention to
Relieve symptoms or decrease symptom severity
Improve functional health status and health-related quality of life
Improve ability to cope with symptoms
Decrease use of health care resources
Avoid harm and be cost-effective

Most FGID intervention studies evaluate the impact of a treatment on the items listed in Table 1, but specific goals can vary widely. Investigators should prioritize their research question(s) pertinent to the specific FGID, develop a hypothesis based on available evidence, and design a study that most effectively answers the research question(s).

In general, the primary question will address whether the study treatment improves FGID symptoms. Consequently, the primary outcome measurement tools must include reporting of the most important symptoms expected to change with the proposed treatment. The secondary questions are best determined by the particular disorder, that is, its specific symptoms and the mechanism of action of the treatment. Pathophysiological factors, while important explanatory parameters, should be considered secondary rather than primary end points.

Defining the Target Condition

Patient Population

A screening log of key variables is mandatory in order for readers to judge the generalizability of the results. The log should include demographic (eg, age, sex, and race) and clinical variables (eg, disease severity, symptom duration, prior treatments for the condition, and the use of concurrent medications) for patients entered and excluded, with reasons for exclusion. Explicit inclusion and exclusion criteria are mandatory for all studies. Most treatment trials in FGIDs have required a minimum severity level for specific symptoms thought to be typical of the condition. Balanced consideration for the potential mechanism of action of the drug must also be given when selecting the study population.

It is advisable to include as broad a spectrum of patients as possible, defined by the Rome- specific FGID criteria. Restricting or modifying the study population must be justified. The EMA requests that early drug development programs include sufficient numbers of both men and women to permit assessment of safety and efficacy for both sexes. The FDA also supports engagement of subjects of different racial backgrounds.[2,3]

Inclusion Criteria

The minimum screening for eligibility should be specified and should adhere to current guidelines. The Rome classification of FGIDs is currently the most comprehensive and well-established diagnostic system, and its use ensures a sufficient degree of standardization of study participants across centers and cultural settings, and allows further exploration for differences in treatment response.

Exclusion Criteria/Appropriate Rule Outs

Important confounding factors to consider for possible exclusion criteria are psychological comorbidities, sociocultural perspectives, and biological variations. Psychological comorbidities are often thought to be predictors of poor response to treatment, but this has not been proven.[4] Other psychologically related influences include the placebo and nocebo effects (see section on placebo and nocebo), and future studies may wish to consider designs that could measure the subject's proneness to these effects.

Managing Functional Gastrointestinal Disorders Overlap, Comorbidities, and Disease Modifiers

Overlap disorders, potential disease modifiers, and important comorbidities that might affect treatment response should be assessed and explored. The overlap of FGIDs with other FGIDs and with somatic and psychiatric disorders is a challenge for clinical trail design. First, the accuracy of the FGID diagnosis may be questioned and it is possible that a treatment might improve the symptoms of one disorder while symptoms of the other worsen. Second, the presence of a comorbidity may be associated with increased symptom severity, greater impact on health-related quality of life (HRQOL), and greater psychological distress—all of which could modify the response to treatment. Third, underlying motility or sensory disorders in different parts of the GI tract may interact in ways that could affect the response to specific treatments. The committee recommends that, in most situations, patients with overlapping conditions be included in the trial and the presence of comorbid conditions should be documented.

Role of Biomarkers in Defining Study Population

Continuing research is needed to identify biomarkers that attempt to elucidate disease mechanisms and may facilitate assessment of efficacy of treatments in FGID studies. A biomarker is an indicator of a physiological or pathological state that can be objectively measured and evaluated, in contrast to PROs, which are measured using questionnaires that capture patient perceptions of their illness.[5] A valid and reliable biomarker should optimally distinguish patients with a known clinical syndrome from other conditions, and do so with a high degree of sensitivity and specificity. It may also have predictive value, in that its presence could potentially predict natural history and/or response to specific therapies.[5] While they are not suitable as surrogate end points at this time, they can be used to stratify patients. However, at present, very few biomarkers have been identified that have sufficient sensitivity and specificity.