

A Comprehensive Review of Randomized Placebo-Controlled Pharmacological Clinical Trials in Children With Functional Abdominal Pain Disorders

*Miguel Saps, *Harman S. Biring, *Cenk K. Pusatcioglu, †Stijn Mintjens, and *Damian Rzeznikewicz

ABSTRACT

Objectives: Abdominal pain–predominant functional gastrointestinal disorders (AP-FGIDs) are the most common cause of consultation to pediatric gastroenterology; however, no medications have been approved to treat this group of disorders in children. The Food and Drug Administration have published recommendations for clinical trials on AP-FGIDs in adults but not in children. The lack of methodological guidelines and accepted primary endpoints for clinical trials in children hampers the progress of the field, making the approval of new medications difficult. A necessary first step to determine the feasibility of clinical trials in children and provide recommendations on the best design for future trials is to review the methods, ability to recruit, attrition rate, and results of previous clinical trials. We designed a comprehensive review of pharmacological clinical trials in AP-FGIDs in children focused on study design.

Methods: Study eligibility was randomized controlled trials (RCTs) evaluating the efficacy of pharmacological interventions compared with that of placebo in children and adolescents with AP-FGIDs.

Results: There is no evidence to support the use of most commonly used drugs in children. Only 7 pharmacological RCTs on AP-FGIDs in children were found. Most studies were single center based and had a small sample size. The methods and outcomes were heterogeneous. Primary endpoints varied widely among studies. Many of the RCTs did not show a consistently significant benefit of the drug over placebo in some or all of the outcomes. We found a considerable risk of bias in most studies. None of the studies have considered minimal clinically important differences in their selection of primary endpoints.

Conclusions: Few randomized clinical trials have been conducted. Most studies have methodological limitations and small sample size. There is an urgent need for well-designed randomized clinical trials using age-appropriate validated outcome measures.

Key Words: abdominal pain, children, clinical trials, functional gastrointestinal disorder

(*JPGN* 2015;60: 645–653)

Received July 7, 2014; accepted January 9, 2015.

From the *Division of Pediatric Gastroenterology, Hepatology, and Nutrition, Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, IL, and the †Department of Pediatrics, Emma Kinderziekenhuis, Academic Medical Centre, Amsterdam, The Netherlands.

Address correspondence and reprint requests to Miguel Saps, MD, Department of Pediatrics, Division of Gastroenterology, Hepatology, and Nutrition, Ann and Robert H. Lurie Children's Hospital of Chicago, 225 East Chicago Ave, Box 65, Chicago, IL 60611-2605, USA (e-mail: MSaps@luriechildrens.org).

The authors report no conflicts of interest.

Copyright © 2015 by European Society for Pediatric Gastroenterology, Hepatology, and Nutrition and North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition

DOI: 10.1097/MPG.0000000000000718

Abdominal pain (AP) affects 38% of school children weekly (1). A subset of all of the children who report AP meets Rome III criteria for an AP-predominant functional gastrointestinal disorder (FGID) (2). FGIDs are the most common cause of consultation to pediatric gastroenterology. Despite the high prevalence of AP-FGIDs, no pharmacological agents are approved for the treatment of these disorders in children in the United States and there is little scientific evidence to support the indication of the most commonly used treatments. This disconnect has long been perceived as an important problem for clinicians and researchers alike. Reassurance and education, key aspects of the consultation in children with AP-FGIDs, are difficult when no scientific data are available. Evidence-based data allow the clinician to establish the patient's prognosis, make an informed decision on when to consider a treatment a failure, and provide an armamentarium of proven treatments in cases of unsatisfactory outcomes.

A key obstacle in designing a pediatric clinical trial on AP-FGIDs is the absence of universally accepted age-appropriate research guidelines. There are no proven biological markers to assess the progress of patients with AP-FGIDs. In the absence of objective means, the evaluation of clinical progress is based on patient-reported outcomes (PROs). In 1999, the Rome committee published a set of guidelines on the design of clinical trials on FGIDs in children (3). The guidelines recommended the use of binary global endpoints of satisfactory relief of symptoms and satisfaction with treatment to establish the clinical benefit of pharmacological treatments. Global binary endpoints have been used for several years in studies in children and adults. Most of the drugs for irritable bowel syndrome (IBS) have been approved based on trials that used global endpoints (4,5). In 2006, the Food and Drug Administration recommended against the use of the historical PROs for IBS because of the lack of a conceptual framework to support them (6). More recently, the Food and Drug Administration published recommendations for clinical trials in adults with IBS that included a new set of clinical endpoints (AP and changes in bowel movements) to substitute the previously used PROs (7). It is unclear whether the new clinical endpoints recommended for adults are applicable for children, superior to the previously used clinical endpoints, or the optimal endpoints that should be used in trials in children with FGIDs. In 2012, the European Medicines Agency (EMA) published their recommendations on the design of clinical trials in adults with IBS (8). The EMA recommendations include a section on children. In this section, the EMA states that “separate trials have to be conducted in children” and that “extrapolation from adults to children—even to adolescents appears to be questionable” and encourages the “development of outcome measures for IBS in children.” In 2013, the Rome IV committee, at the request of the EMA, nominated a subcommittee to establish guidelines and clinical endpoints for clinical trials for IBS in children.

The first step in determining the optimal clinical endpoints for AP-FGIDs in children is to critically and extensively review the existing pediatric literature. We have designed a comprehensive review of pharmacological clinical trials in AP-FGIDs in children focused on study design.

METHODS

Search Strategy

We systematically searched EMBASE, MEDLINE/PubMed, PsycINFO, CINAHL, and the Cochrane Library from inception to October 2013. English- and Spanish-language randomized controlled trials (RCTs) on FGIDs associated with AP in children and adolescents were reviewed. Studies were identified with the term AP (as a medical subject heading). This was combined using the set operator AND with studies identified with the following terms: diet therapy, drug therapy, psychology, surgery, therapy, and etiology (as both medical subject heading and free-text terms). Age was limited to 2 to 18 years old. Study design was meta-analysis or RCT. In addition, literature and systematic reviews, as well as other potentially relevant reference lists, were screened for additional studies of interest. Gray literature such as proceedings from scientific meetings, abstracts, and journal editorials were not included in our search strategy. Two reviewers (D.R., H.S.B.) independently screened the aforementioned databases and the relevant literature for studies. Whenever a conflict ensued during the search, a third reviewer (M.S.) was consulted for conflict resolution.

Eligibility Criteria

We included all of the RCTs that evaluated the effectiveness of pharmacological interventions compared with that of placebo in children and adolescents with functional dyspepsia, IBS, functional AP, and abdominal migraine.

Data Extraction

The data extraction was conducted independently by 2 reviewers (DR, HB). Once it was completed, the results were collated and a third reviewer (MS) was consulted for consensus on data for which inclusion was uncertain. The following data were extracted from each study: inclusion criteria, primary outcomes, study setting, sample size, aims, duration, study design, drug dosage, compliance, and adverse effects.

Risk of Bias

In order to determine the risk of bias of the studies included in the present review, 2 independent reviewers (D.R., H.S.B.) used the Cochrane Collaboration's tool for assessing risk of bias, which evaluated studies based on the appropriateness of their sequence generation, allocation concealment, blinding of personnel and outcome assessors, incomplete outcome data and methodology of addressing incomplete data, and selective outcome reporting (9). We used a stringent criterion: if a certain study did not address one of these indicators or there was not enough information to determine whether the methods used by a study were satisfactory, that domain was judged as "unclear" and scored as 0 regardless of whether the item could have been addressed in the protocol. To minimize the risk of bias in the judgment of this item, MS was excluded from the scoring because he was the author of one of the publications to be evaluated. In order to establish whether some items in the scoring were part of the protocol but not reported in the publication, we attempted to contact the principal investigators of each trial.

RESULTS

The search strategy yielded 5889 studies (Fig. 1). After title and abstract review was conducted, 5781 studies were excluded based on multiple criteria; an adult sample was found to be the main criterion for exclusion by both reviewers. After excluding 7 of the remaining 108 titles because of languages other than English or Spanish, 101 full texts were further screened for inclusion and relevant references lists were assessed, resulting in 24 RCTs (23.7%), which examined interventions on children and adolescents with AP-FGIDs. Lastly, after excluding nonpharmaceutical interventions such as behavioral, probiotic, and dietary interventions, 7 RCTs were included in the present review (Table 1). A total of 325 children and adolescents were recruited to the studies. Five of the 7 RCTs had a parallel-group design (10–14), whereas 2 (15,16) RCTs had a crossover design (none of them included a washout period). All of the studies were double-blinded.

Study Setting

All of the 7 trials were conducted in tertiary care centers or pediatric gastroenterology clinics. Two studies were multicenter (12,14). Five studies were conducted in the United States and the rest in other countries (the United Kingdom and Iran) (13,16).

Inclusion Criteria

There was a wide variation in inclusion criteria throughout the 7 studies. Age range varied among studies, but all of the studies included children of at least 5 years of age (Table 1). Some trials in the review were conducted in patients with a specific AP-FGID

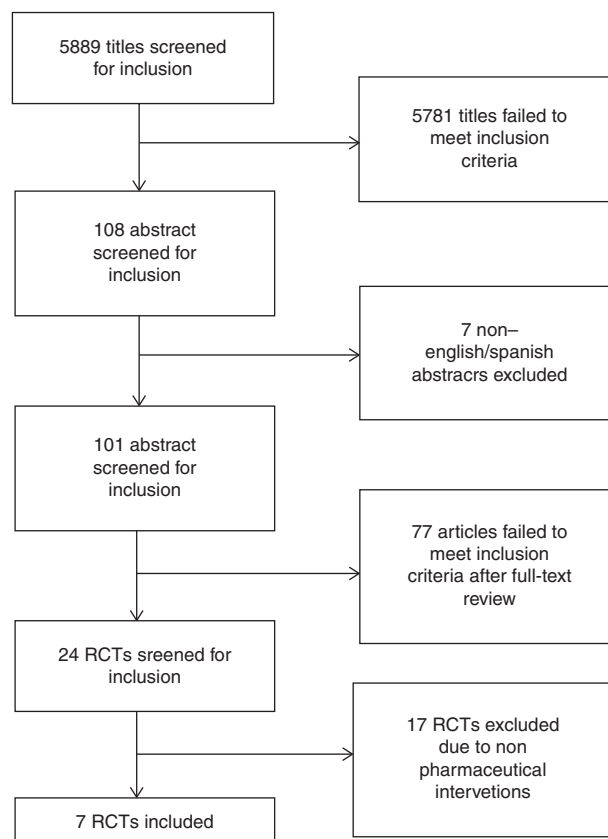


FIGURE 1. Search strategy. RCT = randomized controlled trial.

TABLE 1. Study population

Study	Intervention	Inclusion criteria	Exclusion criteria *	Age, mean years, female, %	Attrition, %	Adverse effects
Bahar et al (10)	Amitriptyline	Rome II criteria for IBS Age: 12–18 y	Prolonged QT	14.7, 73	5.7	None
Collins and Lin (11)	Rifaximin	Rome II criteria for IBS, FD, FAP, abdominal migraine Age: 8–18 y	History of antibiotics or probiotics within 2 mo	12.7, 72	6.7	Intervention—1 (2.0%), abdominal pain (unknown if adverse effect)
Kline et al (12)	Peppermint oil	Manning or Rome I criteria for IBS Age: 8–17 y	Present use of medications for the treatment of IBS or other medications that could affect abdominal symptoms <i>Helicobacter pylori</i> serology, radiography, thyroid function tests	12.0, 60	16.0	None
Sadeghian et al (13)	Cyproheptadine	Rome II criteria for FAP without abdominal migraine or constipation Age: 4–16 y	Clinically significant abnormality on electrocardiogram >3 nighttime awakenings for abdominal pain Abdominal migraine History of constipation, diarrhea, or vomiting Pain relieved with lactose-free diet trial Hypersensitivity to cyproheptadine	7.5, 59	19.4	Intervention—1 (5.5%), hyperactive airway
Saps et al (14)	Amitriptyline	Rome II criteria for FAP, FD, IBS	Abnormal electrocardiogram, tissue transglutaminase, urinalysis, blood count, erythrocyte sedimentation rate, albumin, pancreatic and liver enzymes, or stool examination Positive lactose breath test	12.7, 73	7.8	Intervention—2 (4.3%), mild adverse events: fatigue, rash, headache
See et al (15)	Famotidine	Age: 8–17 y Apley's criteria for RAP without IBS or constipation Dyspeptic symptoms	Abnormal complete blood count, erythrocyte sedimentation rate, pancreatic and liver enzymes, stool examination Diarrhea, constipation, fecal soiling, or bloody stools	10.5, 48	0	Placebo—1 (2.3%), dizziness Not provided
Symon and Russell (16)	Pizotifen syrup	Age: 5–18 y Abdominal migraine for at least 6 mo Age: 5–13 y	Regular acid-suppressive therapy Abnormal hematology profile, urinalysis, erythrocyte sedimentation rate	Not provided	12.5	Intervention—1 (7.1%), drowsiness; greater weight gain (1.25 kg (intervention) versus 0.38 kg (placebo), <i>P</i> = 0.039) Placebo—1 (7.1%), increased appetite

FAP = functional abdominal pain; FD = functional dyspepsia; GI = gastrointestinal; IBS = irritable bowel syndrome; RAP = recurrent abdominal pain.

* All of the studies excluded children with chronic conditions, an organic GI disease, psychological or developmental problems, evidence of growth failure, and abnormal baseline laboratory results.

diagnosis, whereas others included patients with all or most AP-FGIDs. Different diagnostic criteria were used with studies using Apley, Manning, Rome II, Rome III, or variations of these criteria. Three studies were conducted on children with various Rome II criteria diagnoses of AP-FGIDs, whereas the rest were conducted on children with a single diagnosis (IBS, functional dyspepsia, abdominal migraine) (Table 1).

Exclusion Criteria

All of the clinical trials excluded children with other chronic conditions, an organic gastrointestinal disease, psychological or developmental problems, evidence of growth failure, and abnormal baseline laboratory results (Table 1). Studies required various laboratory workup as inclusion criteria, including normal laboratory workup (complete blood count, erythrocyte sedimentation rate, albumin level, pancreatic and liver enzyme levels, urinalysis, stool examination for occult blood and ova and parasites) (12,14,15); 1 study required *Helicobacter pylori* serology, radiography, and thyroid function tests (12), whereas only 1 study requested tissue transglutaminase levels to rule out celiac disease (14) and 2 studies required excluding lactose intolerance by breath hydrogen test or exclusion diet trial (13,14).

Study Duration

The duration of the interventions of the 7 studies varied greatly, ranging from 10 days (11) to 16 weeks (Table 2) (16). Only 2 studies had a run-in period (10,14) ranging from 1 to 2 weeks. The study by Bahar et al was the only one to assess patients 3 weeks after withdrawal of treatment (10).

Sample Size, Rate of Recruitment, and Attrition

Sample size, rate of recruitment, and attrition varied widely between the 7 studies, ranging from 14 (16) to 90 patients (14). Five studies had a sample with female predominance, whereas 1 study did not provide sex information (16) (Table 1). Only 3 studies used statistical power calculations to assess their required sample size (11,14,15). In order to establish the “ability to recruit” children for a clinical trial, we divided the length of time of recruitment by the number of patients enrolled. Rate of recruitment constitutes an approximation because not all of the studies provided the number of months, with some of them providing only the years of recruitment. We decided to provide the information knowing that it may not be completely accurate because recruitment has been found to be a problem in placebo-controlled clinical trials in children. Four of the 7 studies provided information on length of recruitment. See et al enrolled 25 children in a 6-month period (4 per month) (15). Sadeghian et al recruited 36 children in 14 months (2.6 per month) (13), and Saps et al had the longest recruitment period, enrolling 90 children in 44 months (2 per month) (14). Bahar et al recruited 33 children between 2002 and 2005 (8 per year) (10). One study did not sustain any withdrawals, with all of the patients completing the study (15). The rest of the studies had attrition that varied from 6.7% to 19.4%.

Study Aims

All of the studies provided background rationale that the intervention being investigated is associated with pain reduction (10,12–15). Five of the 7 studies presented a study aim to examine the efficacy or benefit of the intervention compared with that of the placebo. The studies by Collins and Lin and Sadeghian et al stated

hypotheses that the intervention would reduce children’s symptoms compared with placebo given to children (11,13). Symon and Russell provided a broader aim by evaluating whether their intervention was useful as a “prophylactic drug” (16).

Outcome Measures

All of the studies used PROs to assess primary and secondary endpoints that varied among studies and included global assessment measures and/or pain reduction. The primary outcome measure in the RCT by Bahar et al was quality of life (QOL), which was self-reported by children (10). A successful outcome was defined as 15% improvement in overall QOL measured by IBS-QOL questionnaire, an adult tool that has not been validated in children (17). The authors adapted the tool by omitting questions on sexual activity. They also assessed multiple secondary outcomes with questionnaires validated in adult patients. Outcome assessments were conducted at 2, 6, 10, and 13 weeks.

Two questions recommended by the Rome II consensus for the design of clinical trials (3) were selected to assess the primary outcome in the multicenter study by Saps et al: a question on satisfactory relief—“Overall how do you feel your problem is?” (“better,” “same,” or “worse”); and a question on satisfaction with treatment—“How did the medication relieve your pain?” (“excellent,” “good,” “fair,” “poor,” or “failed”) (14). The answers to these questions were analyzed in a binary fashion: better and same versus worse and excellent and good versus fair, poor, or failed, and self-reported by children. The study analyzed multiple secondary outcome measures and provided daily diary for the patients to assess daily symptoms including a visual analog–Likert scale ranging from 0 to 100 mm (18).

Collins and Lin assessed frequency and severity of individual gastrointestinal symptoms (bloating, excess gas, incomplete evacuation, AP, diarrhea, constipation, urgency, mucus, straining, fecal incontinence) through a visual analog scale (0–10) that was completed by parents (11). The questionnaire also included 8 multiple-choice questions on the characteristics of their AP (location, frequency, duration) and its effect on daily activities. Participants in the study were also asked to rate their overall symptom improvement as a percentage (100%—complete improvement). Patients underwent a lactulose breath test before and after the trial to assess the effect of the treatment on small intestinal bacterial overgrowth.

The trial conducted by Sadeghian et al did not use validated questionnaires to assess the effect of cyproheptadine on AP (13). Primary outcomes consisted of parents’ and children’s reports of frequency and intensity of AP compared with baseline and both child’s and parental reports of pain progress. Both children’s and parental reports of AP frequency and intensity were measured at 1 and 2 weeks by a 6-point scale (1, completely resolved to 6, worse). For pain progress, children reported using a 4-point scale (1, no pain to 4, worse), whereas parental report was a binary outcome (whether the treatment was satisfactory or not).

The trial by See et al assessed the outcomes of their clinical trial through 2 global primary endpoints that were self-reported by children: quantitative overall AP score and a binary global assessment question (15). The quantitative score, termed AP score, is scored as the sum of 3 subscores measuring pain frequency (pain score—ranging from no pain to multiple times per day), severity (severity score—measured through a validated facial scale with 9 faces, each of them with a corresponding score) (19), and the peptic index score (comprising the presence of nausea, vomiting, chest pain, epigastric pain or tenderness, decreased appetite or weight loss, relation to meals, nocturnal awakening, and pain on waking up in the morning, each of them scored as 1 point). A

TABLE 2. Study intervention characteristics and outcomes

Study	Intervention, dose	Duration of intervention	Study design	Primary outcomes	Results	Positive/negative	Secondary outcomes
Bahar et al (10)	Amitriptyline (n = 16) versus placebo (n = 17) Dose: 10 mg/day (weight 30–50 kg); 20 mg/day (weight 50–80 kg); 30 mg/day (weight >80 kg)	8 wk (run-in period: 2 wk; washout period: 3 wk)	Parallel-group design	QOL	Improvement in overall QOL score at wk 6, 10, and 13 (PP)—intervention versus placebo: 1. wk 6: 127.6 versus 132.0 ($P = 0.019$); 2. wk 10: 128.0 versus 129.4 ($P = 0.004$); 3. wk 13: 126.2 versus 129.8 ($P = 0.013$)	Positive	1. IBS-related symptoms; 2. Pain intensity and frequency (VAS); 3. Interference with daily activities (Likert-like scale); 4. QOL sub scale—dysphoria; (a) Interference with activity; (b) Body image; (c) Health concern; (d) Food avoidance; (e) Social reaction; (f) Relations; (g) Weight gain
Collins and Lin (11)	Rifaximin (n = 49) versus placebo (n = 26) Dose: 550 mg 3 times daily	10 days	Parallel-group design	1. Individual GI symptoms 2. Overall symptom improvement	1. Nonsignificant (data not shown) ITT 2. Nonsignificant (data not shown) ITT	1. Negative 2. Negative	Small intestine bacterial overgrowth (lactulose breath test)
Kline et al (12)	Peppermint oil versus placebo (total n = 42) Dose: 0.2 mL (374 mg) 3 times daily (weight >45 kg); 0.1 mL (187 mg) 3 times daily (weight 30–45 kg)	2 wk	Parallel-group design	1. Symptom improvement 2. Mean pain severity	1. Symptomatic improvement after 2 wk (PP)—intervention: 71%; placebo: 43% ($P < 0.002$) 2. Mean severity of pain (as per daily diary entries) (PP): intervention lower than placebo (data not shown) ($P < 0.03$)	1. Positive 2. Positive	None
Sadeghian et al (13)	Cyproheptadine (n = 14) versus placebo (n = 15) Dose: 0.25–0.5 mg/kg twice daily	2 wk	Parallel-group design	1. Change in pain frequency 2. Change in pain intensity	1. Abdominal pain frequency: “resolved/very improved/improved” (PP)—intervention versus placebo: 1. wk 1: 87% versus 43% ($P = 0.003$); 2. wk 2: 87% versus 36% ($P = 0.002$) 2. Abdominal pain intensity: “resolved/very improved/improved” (PP)—intervention versus placebo: 1. wk 1: 87% versus 36% ($P = 0.001$); 2. wk 2: 87% versus 29% ($P = 0.001$)	1. Positive 2. Positive	None

(Continued)

TABLE 2. (Continued)

Study	Intervention, dose	Duration of intervention	Study design	Primary outcomes	Results	Positive/negative	Secondary outcomes
	Maximum daily: 12 mg for ages 2–6 y; 16 mg for ages 7–14 y			3. Child's assessment of pain progress	3. Children's assessment: "no pain/better" (PP)—intervention versus placebo, 87% versus 36% ($P = 0.005$)	3. Positive	
				4. Parental assessment of pain progress	4. Parent's assessment: "satisfactory versus nonsatisfactory" (PP)—intervention versus placebo, 87% versus 29% ($P = 0.003$)	4. Positive	
Saps et al (14)	Amitriptyline (n = 46) versus placebo (n = 44)	4 wk (run-in period: 1 wk)	Parallel-group design	1. Satisfactory relief	1. Better versus same and worse (ITT)—intervention: 59% versus 34%; placebo: 53% versus 38% ($P = 0.81$)	1. Negative	1. Interference with daily activities; 2. Pain response scores (PR); 3. Depression scores (CDI); 4. Anxiety scores (STAIC); 5. Somatization scores (CSI); 6. Disability scores (FDI); 7. Pain intensity (VAS)
	Dose: 10 mg/day (weight <35 kg); 20 mg/day (weight >35 kg)			2. Satisfaction with treatment	2. Excellent and good versus failed, poor, and fair satisfaction with treatment (ITT)—intervention: 50% versus 43%; placebo: 45% versus 46% ($P = 0.85$)	2. Negative	
See et al (15)	Famotidine (n = 12) versus placebo (n = 13) Dose: 0.5 mg/kg twice daily	6 wk (3 wk followed by crossover for another 3 wk) No washout period	Crossover design	1. Abdominal pain score	1. Mean improvement (PP)—intervention: 3.37; placebo: 1.66 ($P = 0.16$)	1. Negative	Subanalysis of children with dyspeptic symptoms (peptic index score)
				2. Global assessment	2. Better (improvement) (PP)—intervention: 68%; placebo: 12%; McNemar ratio: 5.67 (95% CI 1.64–30.18, statistically significant)	2. Positive	
Symon and Russell (16)	Maximum daily: 40 mg Pizotifen syrup versus placebo (total n = 14) Dose: 0.25 mg twice daily	16 wk (8 wk followed by crossover for another 8 wk) No washout period	Crossover design	1. Days of abdominal pain	1. Mean number of days (PP)—intervention: 4.29; placebo: 12.50 ($P = 0.005$)	1. Positive	None
				2. Index of severity	2. Severity (PP)—intervention: 7.29; placebo: 23.50 ($P = 0.005$)	2. Positive	
				3. Index of misery	3. Misery (PP)—intervention: 25.43; placebo: 81.50 ($P = 0.007$)	3. Positive	

CDI = Children's Depression Inventory; CI = confidence interval; CSI = Children's Somatization Inventory; FDI = Facial Disability Index; GI = gastrointestinal; IBS = irritable bowel syndrome; ITT = intention-to-treat; PP = per-protocol; PRI = Pain-Rating Index; QOL = quality of life; STAIC = State-Train Anxiety Inventory for Children; VAS = visual analog scale.

subanalysis of children with dyspeptic symptoms using the peptic index score (≥ 4) was presented as a secondary outcome.

Symon and Russell evaluated the effect of pizotifen syrup on pain severity through daily diaries and 2 nonvalidated indices: the index of severity and index of misery (16). Questionnaires were completed by children and parents.

The trial by Kline et al used change of overall symptom improvement and mean pain severity as the primary endpoints (12). Patients completed the 15-item Gastrointestinal Symptom Rating Scale (20) on day 1 and again at the end of the trial. Symptoms included AP, change in stool pattern, heartburn, nausea, and vomiting. Daily dairies were used to report changes in the severity of symptoms. Changes in symptoms were ranked 1 to 5 (1, much better to 5, much worse), and severity of pain was ranked from 1 to 5 (1, excellent to 5, much worse).

Analysis of Outcomes (Intention-to-Treat Analysis vs Per-Protocol Analysis)

Only 2 of the RCTs (11,14) analyzed their outcomes based on intention-to-treat analysis (proportion of patients achieving improvement as a percentage of the total number of patients randomized). The method of analysis was unclear in the study by Bahar et al (possibly per-protocol [PP] analysis) (10), whereas the remaining RCTs (12,13,15,16) were clearly based on PP analysis (only patients who completed the study were included in the assessment).

Risk of Bias

The analysis of risk of bias showed that none of the studies was free of risk of bias, with only 3 studies scoring 1 point each of a possible 6 with a score of 6 of 6 indicating the lowest risk of bias (Table 3). Only 2 authors responded to our request for additional information pertaining to measures taken to avoid biases such as random sequence generation, allocation concealment, and blinding of participants, personnel, and outcome assessors. The corresponding author from the study by See et al (15) could not verify the exact details of the methodology; however, the corresponding author stated that the pharmacy at Mt. Sinai Medical Center generated the randomization sequence via a random number generator.

Pharmacological Agents

Two studies (10,14) assessed the efficacy of amitriptyline, a tricyclic antidepressant. The rest of the studies assessed the effect of famotidine (15), rifaximin (11), cyproheptadine (16), peppermint oil (12), and pizotifen.

Efficacy

The study by Bahar et al found improvement in overall QOL score (10) in patients with IBS-associated diarrhea. There was a significant improvement in right lower quadrant pain at 6, 10, and 13 weeks (follow-up), but there was no consistent or significant improvement in pain in other areas. There were no significant differences in the rate of interference with schoolwork, sports, or friends, pain relief after defecation, headache, backache, nausea, dizziness, weakness, constipation, presence of mucous in the stool, tenesmus, or pain exacerbation with defecation between the amitriptyline and placebo groups. The multicenter trial of amitriptyline conducted by Saps et al showed no significant difference in satisfaction with treatment or pain relief in intention-to-treat or PP analysis between drug and placebo (14). Children in both groups

TABLE 3. Risk of bias

Study	Sequence generation	Allocation concealment	Blinding of participants	Blinding of personnel and outcome assessors	Incomplete outcome data	Selective outcome reporting	Overall quality	Overall quality after contacting authors
Bahar et al (10)	Unclear	Unclear	Unclear	Unclear	Yes	Unclear	1/6 with 5 "unclear" criteria	Unable to contact
Collins and Lin (11)	Unclear	Unclear	Yes	Yes	Unclear	Unclear	2/6 with 4 "unclear" criteria	Unable to contact
Kline et al (12)	Unclear	Unclear	Unclear	Unclear	Yes	Unclear	1/6 with 5 "unclear" criteria	Unable to contact
Sadeghian et al (13)	Yes	Yes	Yes	Yes	Yes	Unclear	5/6 with 1 "unclear" criterion	Unable to contact
Saps et al (14)	Unclear*	Unclear*	Yes	Yes	Yes	Unclear	3/6 with 3 "unclear" criteria	5/6 with 1 "unclear" criterion
See et al (15)	Unclear*	Unclear*	Yes	Yes	Yes	Unclear	3/6 with 3 "unclear" criteria	5/6 with 1 "unclear" criterion
Symon and Russell (16)	Unclear	Unclear	Unclear	Unclear	Yes	Unclear	1/6 with 5 "unclear" criteria	Unable to contact

Unclear = indicator not addressed in study; yes = indicator addressed by study was satisfactory.

* Modification of initial assessment after contacting the corresponding author for clarification; "sequence generation" and "allocation concealment" were addressed.

experienced a significant decline in pain, interference with daily activities, somatization, and depression scores from baseline, but no difference was found between the groups. The only significant difference between both groups was in anxiety scores that were lower in the amitriptyline group. In the RCT by Collins and Lin, children receiving rifaximin and placebo group had no significant difference in individual symptoms and overall symptom improvement (11). In the study by Sadhegian et al, children receiving cyproheptadine showed a significant benefit in terms of resolution or improvement in AP intensity and frequency at weeks 1 and 2 compared with those receiving placebo (13). There was a significant difference in both self-report and parental report of child improvement between the intervention and placebo groups. In the RCT by See et al, mixed results were reported for the primary endpoints (15). Although a significant benefit in global improvement was found in famotidine compared with that in placebo, the mean AP score was not significantly different between both groups. The RCT on pizotifen by Symon and Russell found that children in the treatment group had significantly fewer days of AP and lower indices of severity and misery compared with those in the placebo group (16). Kline et al found that children in the peppermint oil group had significant reduction in severity of pain compared with those in the placebo group (12).

Adverse Effects

Mild adverse effects were noted in 4 studies, whereas 2 studies did not report any adverse effect and 1 study did not provide information on adverse effects (13,14,16).

DISCUSSION

We have conducted a methodological review of RCTs on pharmacological agents in children with AP-FGIDs. The importance of the present review transcends the enumeration of the various features of each study. The understanding of the ability to recruit in each study setting, methodology, and comparative results when different outcome measures are used is key to establish the feasibility and specific recommendations for the design of RCTs in children. Although the argument could be made that pediatric study guidelines could be based on the same principles that guide adult studies, the results of the present review suggest otherwise. The results of pediatric studies on the efficacy of amitriptyline for the treatment of AP-FGIDs are in contrast with the results of adult studies. Although adult studies on tricyclic antidepressants found a positive effect of this group of drugs in the treatment of IBS (21), a systematic review of the efficacy of antidepressants in children and adolescents with AP-FGIDs by the Cochrane group found no evidence to recommend the use of antidepressants in children (22). This contradicts the results of multiple meta-analyses in adults with IBS (23–26) and the results of an evidence-based position statement by the American College of Gastroenterology that tricyclic antidepressants are effective for adult patients with IBS (27). Conflicting results between adult and pediatric studies can be interpreted as a different effect of the drug in children and adults. We, however, cannot exclude the fact that differences in understanding or relevance attributed to global outcomes in adults and children (children may have a different concept of satisfactory relief or satisfaction with treatment than adults) may explain the opposite results. Answering this question is beyond the scope of the present review, but the likelihood of children and adults requiring different outcome measures cannot be ignored. Establishing primary endpoints to support the efficacy of a particular drug should reflect meaningful changes for each age group. Although adult studies have established minimal clinically important differences (MCIDs)

(minimal clinical change that the patient identifies as meaningful) for trials on IBS (28), no studies have been published in pediatrics (data are being analyzed by our group) and none of the RCTs in our review has considered MCIDs at the time of selection of outcomes. The absence of data on pediatric MCIDs precludes us from making recommendations on primary efficacy endpoints that reflect what is clinically meaningful to children.

Many of the RCTs in the present review did not show a consistently significant benefit of the drug over placebo in some or all of the outcomes. We cannot exclude the fact that these negative results could be at least partially explained by the use of an inclusion criterion that was exclusively based on AP-FGID phenotypes ignoring the fact that children in the sample could have different pretrial expectations of success, dissimilar psychosocial milieu, and various pathophysiological mechanisms. AP-FGIDs are complex and multifactorial disorders. A better characterization of the various factors involved in the pathogenesis and pathophysiology of different groups of children could help future studies. Establishing a more focused inclusion criterion based on specific psychosocial profiles and pathophysiological mechanisms consistent with the mechanism of the drug to be studied may increase the effect size, making it possible to demonstrate a significant difference between drug and placebo with a smaller sample size. The design of studies with a smaller sample would be of particular importance considering the difficulties in recruitment found in our review. All of the studies had a modest monthly ratio of recruitment, with none of the studies been able to recruit >4 children per month. Once patients were enrolled in the study, however, there was a low attrition rate, which suggests that families who accepted to participate in the trials were not discouraged by participating in a study with a placebo arm despite the poor outcomes achieved by many patients in the studies. We found a variable placebo effect, with some studies showing a placebo effect (10) and others a high placebo effect (14). This variation in the placebo effect among studies warrants further investigation into the various factors that may influence the placebo effect. Understanding these factors may allow designing studies comparing patients receiving the active drug with patients with a low likelihood of achieving a good response while receiving placebo, a strategy that may also allow increasing the effect size and decreasing the sample size required to find a significant difference between drug and placebo.

Despite the high prevalence of AP-FGIDs, we found that only 7 RCTs with a total of 325 children were ever published. Moreover, studies were generally small, had heterogeneous and sometimes unclear methods, and were prone to bias. Some of the studies were conducted on drugs that are now rarely used, whereas others used old definitions that are no longer used (recurrent AP) or nonvalidated criteria for diagnosis. Shortcomings in design of the various RCT underscore the need for education and the establishment of detailed guidelines for clinical trials in children. We hope that the present review serves to stimulate researchers to conduct further studies while providing information on preventable shortcomings. Future guidelines should include information on study design, minimum length of trials, inclusion/exclusion criteria, primary efficacy endpoints, MCID, validated outcome measures, and reliability and sensitivity of the outcome measures recommended to be used in each AP-FGID. Optimization of design and reporting is paramount, considering the great need for clinical trials and the important difficulties encountered by most of the authors in recruiting patients for the studies. Despite the small number of RCTs found and the shortcomings encountered in many of the studies, we found some encouraging signs that raise our hopes in the progress of pediatric neurogastroenterology in the near future. Most of the studies and the largest RCTs were conducted in the past 6 years, with 234 of 325 participating patients (72%) having been

recruited in this time period, compared with those in the previous 15 years (28%).

Our review is not devoid of limitations. We have reviewed only the literature in English and Spanish. We cannot exclude the fact that our review did not report other RCTs that were published in other languages. Our review, however, included studies conducted in distant countries such as Iran that were published in English. We have not reviewed data on RCTs conducted without a placebo arm and open-label trials. Considering the high placebo effect found in some of the studies and the nocebo effect found in others that would make the interpretation of studies without a placebo arm difficult, we made the decision to limit our review to studies comparing drug with placebo. In conclusion, our review found only 7 pharmacological RCTs on AP-FGIDs in children. Most of the studies had methodological limitations and a small sample size. The studies used varied methodology of inclusion, assessment, and length. These data do not allow establishing recommendations on the design of clinical trials in AP-FGIDs in children. Studies on MCIDs, validated outcome measures, and clinical endpoints in children are needed. Standardized validated questionnaire banks such as PROMIS (29) may help clinical practice and research.

REFERENCES

- Saps M, Seshadri R, Sztainberg M, et al. A prospective school-based study of abdominal pain and other common somatic complaints in children. *J Pediatr* 2009;154:322–6.
- Rasquin A, Di Lorenzo C, Forbes D, et al. Childhood functional gastrointestinal disorders: child/adolescent. *Gastroenterology* 2006;130:1527–37.
- Veldhuyzen van Zanten SJ, Talley NJ, Bytzer P, et al. Design of treatment trials for functional gastrointestinal disorders. *Gut* 1999;45 (suppl 2):II69–77.
- Krause R, Ameen V, Gordon SH, et al. A randomized, double-blind, placebo-controlled study to assess efficacy and safety of 0.5 mg and 1 mg alosetron in women with severe diarrhea-predominant IBS. *Am J Gastroenterol* 2007;102:1709–19.
- Novick J, Miner P, Krause R, et al. A randomized, double-blind, placebo-controlled trial of tegaserod in female patients suffering from irritable bowel syndrome with constipation. *Aliment Pharmacol Ther* 2002;16:1877–88.
- Food and Drug Administration. Guidance for industry. Patient-reported outcome measures: use in medical product development to support labeling claims. <http://www.fda.gov/ohrms/dockets/98fr/06d-0044-gdl/0001.pdf>. Published 2006. Accessed May 27, 2014.
- Food and Drug Administration. Guidance for industry. Irritable bowel syndrome—clinical evaluation of drugs for treatment. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM205269.pdf>. Published 2012. Accessed May 27, 2014.
- European Medicines Agency. Guideline on the evaluation of medicinal products for the treatment of irritable bowel syndrome. http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2013/07/WC500146176.pdf. Published 2013. Accessed May 27, 2014.
- Higgins JP, Altman DG, Gotzsche PC, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
- Bahar RJ, Collins BS, Steinmetz B, et al. Double-blind placebo-controlled trial of amitriptyline for the treatment of irritable bowel syndrome in adolescents. *J Pediatr* 2008;152:685–9.
- Collins BS, Lin HC. Double-blind, placebo-controlled antibiotic treatment study of small intestinal bacterial overgrowth in children with chronic abdominal pain. *J Pediatr Gastroenterol Nutr* 2011;52:382–6.
- Kline RM, Kline JJ, Di Palma J, et al. Enteric-coated, pH-dependent peppermint oil capsules for the treatment of irritable bowel syndrome in children. *J Pediatr* 2001;138:125–8.
- Sadeghian M, Farahmand F, Fallahi GH, et al. Cyproheptadine for the treatment of functional abdominal pain in childhood: a double-blinded randomized placebo-controlled trial. *Minerva Pediatr* 2008;60:1367–74.
- Saps M, Youssef N, Miranda A, et al. Multicenter, randomized, placebo-controlled trial of amitriptyline in children with functional gastrointestinal disorders. *Gastroenterology* 2009;137:1261–9.
- See MC, Birnbaum AH, Schechter CB, et al. Double-blind, placebo-controlled trial of famotidine in children with abdominal pain and dyspepsia: global and quantitative assessment. *Dig Dis Sci* 2001;46:985–92.
- Symon DN, Russell G. Double blind placebo controlled trial of pizotifen syrup in the treatment of abdominal migraine. *Arch Dis Child* 1995;72:48–50.
- Patrick DL, Drossman DA, Frederick IO, et al. Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. *Dig Dis Sci* 1998;43:400–11.
- Tesler MD, Savedra MC, Holzemer WL, et al. The word-graphic rating scale as a measure of children's and adolescents' pain intensity. *Res Nurs Health* 1991;14:361–71.
- McGrath PA. The multidimensional assessment and management of recurrent pain syndromes in children. *Behav Res Ther* 1987;25:251–62.
- Svedlund J, Sjodin I, Dotevall G. GRS—a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. *Dig Dis Sci* 1988;33:129–34.
- Drossman DA, Toner BB, Whitehead WE, et al. Cognitive-behavioral therapy versus education and desipramine versus placebo for moderate to severe functional bowel disorders. *Gastroenterology* 2003;125:19–31.
- Kaminski A, Kamper A, Thaler K, et al. Antidepressants for the treatment of abdominal pain-related functional gastrointestinal disorders in children and adolescents. *Cochrane Database Syst Rev* (7):2011:CD008013.
- Chao GQ, Zhang S. A meta-analysis of the therapeutic effects of amitriptyline for treating irritable bowel syndrome. *Intern Med* 2013;52:419–24.
- Ruepert L, Quartero AO, de Wit NJ, et al. Bulking agents, antispasmodics and antidepressants for the treatment of irritable bowel syndrome. *Cochrane Database Syst Rev* 2011;8:CD003460.
- Rahimi R, Nikfar S, Rezaie A, et al. Efficacy of tricyclic antidepressants in irritable bowel syndrome: a meta-analysis. *World J Gastroenterol* 2009;15:1548–53.
- Ford AC, Talley NJ, Schoenfeld PS, et al. Efficacy of antidepressants and psychological therapies in irritable bowel syndrome: systematic review and meta-analysis. *Gut* 2009;58:367–78.
- American College of Gastroenterology Task Force on Irritable Bowel Syndrome Brandt LJ, Chey WD, Foxx-Orenstein AE, et al. . An evidence-based position statement on the management of irritable bowel syndrome. *Am J Gastroenterol* 2009;104 (suppl 1):S1–35.
- Spiegel B, Camilleri M, Bolus R, et al. Psychometric evaluation of patient-reported outcomes in irritable bowel syndrome randomized controlled trials: a Rome Foundation report. *Gastroenterology* 2009;137:1944–53e1–3.
- PROMIS. Dynamic tools to measure health outcomes from the patient perspective. <http://www.nihpromis.org>. Accessed October 10, 2014.