REVIEW ARTICLE: ENDPOINTS USED IN FUNCTIONAL DYSPEPSIA DRUG THERAPY TRIALS

Daphne Ang, Nicholas J Talley, Magnus Simrén, Pieter Janssen, Guy E Boeckxstaens, Jan Tack

Alimentary Pharmacology & Therapeutics

REVIEW ARTICLE: ENDPOINTS USED IN FUNCTIONAL DYSPEPSIA DRUG THERAPY TRIALS

Journal:	*Alimentary Pharmacology & Therapeutics*
Manuscript ID:	APT-0544-2010.R1
Wiley - Manuscript type:	Review Article
Date Submitted by the Author:	11-Dec-2010
Complete List of Authors:	Ang, Daphne; Changi General Hospital, Gastroenterology Talley, Nicholas; University of Newcastle, Faculty of Health Simrén, Magnus; Section of Gastroenterology and Hepatology, Dept of Internal Medicine Janssen, Pieter; KULeuven Boeckxstaens, Guy; University of Leuven, Pathophysiology Tack, Jan; University Hospital, Center for Gastroenterological Research
Keywords:	Functional dyspepsia < Disease-based, Functional GI diseases < Disease-based, Stomach and duodenum < Organ-based, Symptom score or index < Topics

REVIEW ARTICLE: ENDPOINTS USED IN FUNCTIONAL DYSPEPSIA DRUG THERAPY TRIALS

D. Ang[1], N.J. Talley[2], M. Simren[3], P. Janssen[1], G. Boeckxstaens[1], J. Tack[1]

[1]Center for Gastroenterological Research, K.U.Leuven, Belgium.
[2]Department of Internal Medicine, Mayo Clinic Jacksonville
[3]Department of Internal Medicine, Sahlgrenska University Hospital, Göteborg, Sweden,

Correspondence to: Jan Tack, M.D., Ph.D.
Department of Pathophysiology,
Division of Gastroenterology
University Hospital Gasthuisberg
Herestraat 49, B-3000 Leuven, Belgium.
Phone: +32 16 345751
Fax: + 32 16 345939
E-mail: Jan.Tack@med.kuleuven.ac.be.

Key words: Functional dyspepsia, endpoints, clinical trials, patient reported outcomes, quality of life, questionnaires

Conflict of interest: none declared

Concept: J. Tack, NJ Talley

Literature source selection: all authors

First draft: D. Ang

Corrections and text finalisation, all authors.

ABSTRACT

Background: How to measure patient reported outcomes (PRO) in treatment trials for functional gastrointestinal disorders is a matter of controversy. This review focuses on instruments and endpoints that have been used to evaluate the efficacy of therapeutic agents in functional dyspepsia (FD) trials, also considering the newly defined Rome III FD criteria. Methods: A Medline search was conducted to identify relevant studies pertaining to FD treatment, with particular emphasis on the studies to date which have used validated outcome measures. Results: Currently available outcome measures are heterogeneous across studies. They include global binary endpoints, analog or categorical scoring scales, uni- or multi-dimensional disease specific questionnaires, global outcome evaluations, and quality of life questionnaires. Across the available outcome measures, substantial heterogeneity is found, not only in the type of endpoint measure, but also in the number and types of symptoms that are considered to be part of the FD symptom complex. Especially based on content validity, none of the existing questionnaires or endpoints can be considered sufficiently validated to be recommended unequivocally as the primary outcome measure for FD trials according to the Rome III criteria. On the other hand, existing well-validated multi-dimensional questionnaires that include many non-FD symptoms can be narrowed down to evaluate only the cardinal symptoms according to Rome III. In conclusion, there is an urgent need to develop Rome III-based PROs for functional dyspepsia. Well-validated multi-dimensional questionnaires may serve as a

guidance for this purpose, and could also be considered for use in ongoing clinical trials.

INTRODUCTION

Functional dyspepsia (FD), is one of the most prevalent functional gastrointestinal disorders (FGIDs). Over the last 20 years, the definition of FD has undergone major changes from the 1988 working party[1] to the consecutive Rome consensus documents[2-4], in line with changing understanding of the pathophysiological basis of this disorder. Based on Rome III criteria[4], FD is defined as the presence of symptoms thought to originate in the gastroduodenal region (early satiation, postprandial fullness, epigastric pain or burning), in the absence of any organic, systemic or metabolic disease that is likely to explain the symptoms. FD is further subdivided into 2 diagnostic categories of meal induced dyspeptic symptoms (postprandial distress syndrome [PDS], characterized by postprandial fullness and early satiation) and epigastric pain syndrome ([EPS], characterized by epigastric pain and burning). Whereas the previous Rome II definition[3] for FD excluded patients with predominant heartburn and was unclear on non-predominant heartburn, the Rome III definition states that heartburn is not a gastroduodenal symptom, although it often occurs simultaneously with FD symptoms.[4] Similarly, although there is frequent overlap with the irritable bowel syndrome, the symptom pattern of both entities is distinct.

FD is one of the most important categories of the FGID, in view of its prevalence and impact on the general population.[5-8] At present, there is no treatment with established efficacy for FD.[1] The Food and Drug Administration (FDA) guidance released in February 2006 provides recommendations for the

use of validated instruments to assess treatment outcomes, and describes the proper development and psychometric validation of patient reported outcomes (PRO) before endorsing a clinical product.[9] The Rome III committee[10] has also provided guidelines for clinical trial design in FGIDs, with a similar emphasis on individual patient assessment and the use of validated outcome measures. The aim of the present study is to evaluate the currently available endpoints for FD drug development, in line with the Rome III definitions, the published FDA guidelines and Rome III clinical trials recommendations.[4,9,10]

MATERIALS AND METHODS

Identification of relevant studies

To identify relevant studies, both computerized (Medline) and manual searches were performed, using the cited references of the retrieved articles. MeSH and free-text terms for FD therapies were combined with the terms dyspepsia, clinical trials, symptom assessment, questionnaires, patient reported measures and randomized for searches conducted for the time period between January 1979 and December 2008. FD therapeutic trials were retrieved and studied, with emphasis on the method used for evaluating outcome measures. Article reference lists were examined for relevant papers. Analysis of both full papers and abstracts was conducted, with particular emphasis on outcome measures. The literature search was performed independently by 2 of the authors, and their retrievals were merged.

Inclusion criteria

The following criteria were used to select trials for analysis: (i) randomized double blind controlled trials (RCT), (ii) parallel or single cross-over trial designs, (iii) adult patients with FD, (iv) baseline gastroscopy to exclude structural pathology; (v) comparison of therapy vs. active or placebo control; (vi) clear description of the method of assessing outcome measures for FD symptoms, and (vi) articles in English.

Outcome measures

One of the main difficulties encountered with therapeutic trials in FGIDs, and FD in particular where efficacy has not been established for any treatment, is the lack of objectively measurable outcome measures.[11-14] In FD, symptom patterns do not correlate well with putative pathophysiological mechanisms such as gastric accommodation, sensitivity to distension and gastric emptying rates.[15] In the absence of quantifiable surrogate markers for symptom improvement, evaluation of treatment response has to rely on patient's reporting of symptom intensities.

Several types of outcome measures have been used in FD clinical trials. These can be broadly classified into global outcomes, generic instruments and disease specific instruments.[16] Global outcomes are measured either by a dichotomous binary type response (e.g. yes/no to symptom improvement) or a scoring method which can either be a Likert scale (categorical) or a visual analogue scale (VAS). Likert scales allow easier interpretation for the physician compared to the VAS, with 5-point or 7-point scales providing greater sensitivity than 4-point scales.[17] The VAS and 7-point Likert scales exhibit comparable responsiveness, although the ease of administration and interpretation of the 7-point scale recommend its use in clinical trials.[18] The Likert scale is often incorporated into global scales, such as the overall treatment effect (OTE),[19] generic instruments and disease specific instruments. The OTE uses a graded scale to assess overall symptom improvement or deterioration. The subjective nature of this approach allows the individual to integrate all aspects of his condition into a single treatment

outcome, and is particularly suitable to show deterioration. Unlike generic instruments, disease specific instruments focus on a particular medical condition and are likely to be more useful in quantifying changes in quality of life in FD therapeutic trials.[20] Disease specific instruments in FD can be unidimensional (evaluating gastrointestinal symptoms) or multidimensional (evaluating both gastrointestinal symptoms and other domains such as emotional or social functioning and impact of symptoms on daily activities).

Validity of outcome measures

In 1999, the Rome II Working Group emphasised individual patient assessment as a primary outcome, allowing for the integration of multiple symptoms into a single global endpoint for FGID studies. [21] An alternative method was the use of a disease specific questionnaire to evaluate relevant aspects of the patient's symptoms and disease related quality of life. The Rome III guidelines and the FDA February 2006 guidance provide further recommendations on the proper development and validation of PRO measures.[9,10] Psychometric validation of a symptom based measure incorporates several components, including evidence that the instrument addresses all the patient's symptoms that are indicative of the disorder (content validity); is related to other measures of the same or similar concepts such as symptom improvement and symptom free days (construct validity); produces similar results when re-administered to patients whose health status has not changed (reliability); detects clinically meaningful change in health status when a change has occurred (longitudinal construct validity or

responsiveness); and is associated with changes in score that can be related to clinical indicators that are meaningful to clinicians (predictive or criterion validity). Responsiveness is an important component of validity[22] and forms a major criterion for selecting an optimal measure of outcome for a randomized controlled trial.[23] Patient involvement in the development of PRO measurement is emphasised by the FDA guidelines, and this can be aided by structured interview sessions, focus groups and quality research methods. The outcome measure should have an effective measurement range[24], allowing the instrument to detect changes in outcomes during the course of the clinical trial without the limitations of the ceiling or floor effects.

RESULTS

Description of studies

Amongst the therapeutic trials for FD, 117 studies were initially obtained. After excluding 31 studies which did not satisfy the inclusion criteria, 86 studies were selected for analysis of the outcome measures. The outcome measures that were employed in these studies can be classified broadly under the following categories: (a) binary outcome measures (adequate, satisfactory or sufficient relief), (b) individual symptom assessment with Likert or VAS scales, (c) disease specific questionnaires, (d) global outcome evaluations, and (e) quality of life questionnaires. Table 1 summarizes the types of endpoints used in the FD studies that were selected for the analysis.

A. Adequate relief or satisfactory relief (Binary Outcome) as a primary endpoint

Evaluating the relief of FD symptoms using a binary outcome approach (yes/no response) was first employed in a study which was designed to compare reflux episodes in FD patients who responded to omeprazole with non-responders. [25] Responders were defined as patients who considered themselves to have achieved sufficient relief of symptoms at the end of study period. A significantly higher number of reflux episodes on 24 hour oesophageal pH measurement were detected in omeprazole responders. Based on extensive experience from irritable bowel syndrome (IBS) studies,[26-

[37] adequate relief and satisfactory relief have been the most commonly used outcome measures in treatment trials for FGIDs, although the Rome III committee on design of treatment trials indicated that further validation may be desirable, and that alternative outcome measures such as integrative symptom questionnaires are also attractive.[10]

The endpoint of adequate relief is responsive, reproducible and demonstrates good construct validity,[38,39] and has been used in both IBS[26-33] and FD[40-42] clinical trials. It allows the patient to integrate all relevant symptoms and normalises the assessment to the patient's own internal reference system.[38,39] On the other hand, the endpoint does not reflect the magnitude of improvement needed to reach the endpoint (which depends in part on the baseline severity level) and does not detect worsening of symptoms[39]. Adequate relief was used as a primary endpoint in a phase 2b RCT of alosetron for FD[40]. During the 12 weeks of treatment with alosetron or placebo, patients responded yes or no to the weekly question: "In the past 7 days, have you had adequate relief of your upper abdominal pain or discomfort?" A 1 mg dose of alosetron showed benefit over placebo and 0.5 or 2 mg doses for the 12-week average rate of adequate relief. When stratified by gender, alosetron 1.0 mg showed significant benefit in females ($p=0.03$) in achieving the primary endpoint of adequate relief. Secondary efficacy endpoints were patients' daily self rating of symptom frequency (absence or presence of early satiety, bloating, nausea, belching, pain and postprandial fullness) and severity (on a 5-point Likert scale). When the primary global outcome of adequate relief was compared with the secondary

outcome of individual symptoms, patients with adequate relief of upper abdominal pain or discomfort also had significantly greater reductions in severity of pain, nausea and bloating, and % of days with pain, early satiety, bloating and nausea compared to patients who did not have adequate relief (all $p<0.001$). Thus, the adequate relief endpoint showed good correlation when individual symptoms were analysed, but no further validation in FD has occurred. The adequate relief endpoint was also used in a phase 2 U.S. study with Acotiamide, a muscarinic receptor antagonist with fundus-relaxing properties.[41] During the first 4 weeks of treatment, the number of subjects reporting adequate relief for more than 50% of time was significantly higher with acotiamide 300 mg t.i.d. compared to placebo.[42] This was associated with significant improvement of symptoms of bloating, postprandial nausea, and stomach pain before meals, and with significant improvement of several domains of the Short Form-36 quality of life scale and of the Short form Nepean Dyspepsia Index.[42]

The satisfactory relief endpoint has a construct that is similar to adequate relief, although less extensive validation data are available in the literature for the former.[39] Using satisfactory relief as an endpoint, the beneficial effects of tegaserod in IBS have been demonstrated.[34-37] The binary satisfactory relief endpoint was also used in two phase 3 studies with tegaserod in dysmotility-type FD, and demonstrated significant efficacy in one of these.[43] In the trial that achieved a significantly higher proportion of days with satisfactory relief from tegaserod, patients also had significant improvement of overall and

individual symptom severity measured on a Likert scale, and a significantly better global assessment of change in dyspepsia symptoms.

B. Categorical (Likert) Scales

Likert scales have been widely used both in assessing individual and global symptom severity as well as symptom frequency in numerous FD trials.[44-97] Veldhuyzen and colleagues[93] used a 5-point Likert scale (0=no problem, 4=very severe problem) to score the severity of 8 gastrointestinal symptoms (epigastric pain, burping, heartburn, bloating, flatulence, sour taste, nausea and halitosis). A validation study was performed in non ulcer dyspepsia (NUD) and *Helicobacter pylori* associated gastritis (HPAG). Its reproducibility was demonstrated by comparing symptom scores in the 2 groups of patients before treatment, and responsiveness was demonstrated in both NUD and HPAG patients, where cumulative scores were significantly reduced after therapy. However, the validity of this scale as a measure of improvement was only assessed by comparing changes in symptom scores with changes in the patients' overall health status (deteriorated, stayed the same or improved during the trial) rather than against an established questionnaire. In addition, it combined FD symptoms with symptoms suggestive of GERD, nausea and intestinal dysfunction. Finally, the cumulative symptom score attributed equal weight to all symptoms, although some of them may be related. Nevertheless, the results from this study support the applicability of a Likert-scale in measuring the severity of FD symptoms.

Likert scores have been employed in numerous studies to date (Table 1), including the BOND and OPERA studies[88] which utilised a 4-point Likert scale to evaluate the efficacy of omeprazole in FD. A 7-point Likert scale was used in 2 very similar multicentre, multinational, RCTs: the OCAY[89] and ORCHID studies,[90] which compared the efficacy of *Helicobacter pylori* eradication therapy with that of either omeprazole (OCAY study) or placebo (ORCHID study) in relieving dyspeptic symptoms 12 months after eradication therapy. Using a stringent definition of treatment success (absence of symptoms or the presence of only minimal symptoms (Likert score 1 or 2) during the seven days preceding the final visit) no significant symptomatic benefit of *Helicobacter pylori* eradication therapy in FD was found.[89,90]

In addition, Likert scales have been used for global evaluation. Using a Likert-type scale with 5 grades (symptom-free, markedly improved, moderately improved, not changed, and deteriorated) to measure patients' global assessment of efficacy, a statistically significant benefit was found in a Phase IIb trial for itopride over placebo in the number of patients who reported being symptom-free or with marked improvement.[91] However, this effect was not seen when the same global patient assessment of efficacy was evaluated in 2 phase III trials. [92]

Another method of assessing global outcome is the "overall treatment effect" (OTE) approach.[18] At intervals during or at the completion of treatment, the

patient is asked to decide whether symptoms have remained the same, improved or deteriorated compared to pre-treatment phase, by means of a Likert scale. The advantage of this endpoint is that it closely resembles the way physicians evaluate treatment benefit in clinical practice, but the main disadvantage is the inherent recall of pre-treatment symptom severity which may lead to bias. OTE was used as a secondary outcome measure in the OCAY study, and in a number of studies of acid suppression in FD.[89,94-97] Overall treatment evaluation was used most recently in the tegaserod FD studies, where a weekly global assessment of change was rated on a 7-point Likert scale, and this endpoint generated the most consistent improvements with tegaserod over placebo.[43] OTE was also used in the U.S. and Japanese Acotiamide trials as a secondary or primary endpoint respectively.[42,98] In the U.S. study, benefit using the OTE paralleled the adequate relief outcome result, with significance during the first 4 to 6 weeks.[42] In the Japanese study, significant benefit was found using the OTE for the 100 mg dose in meal-related FD symptoms, and this was associated with a higher rate of disappearance of postprandial fullness and upper abdominal bloating.[98] VAS scores, which may be more difficult to interpret in terms of magnitude of response, have been used in only a few, mainly less recent studies.[99-104]

C. Disease specific questionnaires (Unidimensional)

Integrative symptom questionnaires incorporate the frequency and/or severity of one or a group of symptoms pertaining to the FGID in question both at baseline and at the end of therapy. A review of the questionnaires that have

been used as outcome measures in FD drug therapy trials is pertinent, as they serve an important tool in addressing the effectiveness of pharmaceutical agents. We highlight several instruments that have been developed for assessing outcome measures in FD. Where appropriate, FD trials that have utilised the respective outcome measures are described.

C. 1. Global Overall Symptom scale (GOS)

The Global Overall Symptoms (GOS) scale is a validated outcome measure for dyspepsia treatment trials,[105] adapted from the previously validated 5-point scale.[93] Using a 7-point Likert scale, patients are asked to grade the overall severity of 10 upper gastrointestinal symptoms (epigastric pain; epigastric discomfort; heartburn; acid regurgitation; upper abdominal bloating; excessive belching; nausea; early satiety; postprandial fullness and persistent fullness after a meal) [specific symptom subtypes (SSS)] over a certain retrospective period of time, either 28 days (GOS-28) or 2 days (GOS-2). Validation of the GOS was tested within the CADET-HN study[94] and the Confirmatory Acid Suppression Test (CAST) study.[106] The GOS showed construct validity, reliability and responsiveness.[105] Construct validity was determined using Spearman correlation coefficients, by correlating changes in the GOS to changes in severity of individual and mean symptom scores; Quality of Life in Reflux and Dyspepsia (QoLRAD)[107] overall score and dimensions the Gastrointestinal Symptom Rating Scale (GSRS, see below)[108] overall score and dimensions, and the Reflux Disease Questionnaire (RDQ)[109] overall score and dimensions. Moderate to high Spearman correlation coefficient

values (0.41-0.80) were achieved for each of the above mentioned outcome measures, demonstrating good construct validity. Test-retest reliability, assessed by Intraclass Correlation Coefficient (ICC) was 0.62 for GOS-2 and 0.42 for GOS-28. In addition, there was a positive correlation between change in GOS and change in symptom severity. To date, the GOS has been used as a primary outcome measure in the CADET (Canadian Adult Dyspepsia Empiric Treatment) study programmes as well as in the ENTER trial.[95-97] In this multicentre placebo-controlled trial carried out across Canada, patients with FD of moderate severity (defined as a GOS score ≥4) were randomized to receive esomeprazole or placebo once daily for 8 weeks. For the primary outcome measure of symptom relief (GOS score ≤2 at 8 weeks), no statistically significant difference was found between both groups. However, the GOS combines FD symptoms with symptoms suggestive of GERD or IBS, and the mean score attributes equal weight to all contributing symptoms, which might be considered to be a drawback.

C.2. Gastrointestinal Symptom Rating Score (GSRS)

The GSRS[108] was developed in the early 1980s as an outcome measure for peptic ulcer disease and IBS. It comprises 15 items incorporating 5 symptom clusters (gastroesophageal reflux, abdominal pain, indigestion, diarrhoea and constipation) over the prior 2 weeks. Although it was originally designed to be interview based with a 4-point adjectival scale, it was subsequently modified to be self administered with a 7-point Likert scale. The GSRS has been well validated and shown to be responsive.[109-111] A shortcoming of the GSRS is

that it is not specific for dyspepsia. In addition, it measures only the severity of GI related symptoms and not their impact on quality of life. The GSRS has been used as a primary outcome measure in a study evaluating the efficacy of rebamipide[112] and as a secondary outcome measure in various studies.[88,94-97,113,114]

C.3. Leeds Dyspepsia Questionnaire (LDQ)

The LDQ[115] is administered by an investigator during an interview, and evaluates the frequency and severity of 8 symptoms (epigastric pain, retrosternal pain, regurgitation, nausea, vomiting, belching, early satiety and dysphagia) The scoring system uses the frequency of the first 5 symptoms to determine the presence of dyspepsia, whilst all 8 symptoms are used to access severity of dyspepsia on a scale of 0 (absent) to 5 (most severe). The severity of dyspepsia is assessed as a summary score (range 0-40). The LDQ was shown to be reliable, valid and responsive to change in both primary and secondary care populations in the U.K. Several potential drawbacks of the LDQ include the administration of the questionnaire by a researcher, inclusion of GERD and nausea symptoms, and a complex scoring system.

The psychometric properties of the LDQ were evaluated in 99 primary care patients and 215 hospital referral patients. Physician diagnosis was the gold standard used to validate the ability of the LDQ to detect dyspepsia. The sensitivity and specificity of the LDQ at the primary care setting were 80% and 79%, respectively. In the hospital setting, the LDQ showed a sensitivity of

99% whilst the specificity was 53%. Based on κ statistics, there was moderate to substantial agreement between the LDQ and physician assessment. Test-retest reliability was evaluated in 107 patients by a research nurse during two clinic visits within 4-7 days, yielding a k statistic of 0.83 and internal consistency by Cronbach's α of 0.68. A similar high k statistic of 0.90 was achieved when two different researchers administered the LDQ to 42 patients within 30 minutes. Responsiveness to change was also detected, with the median LDQ score decreasing from 22.5 to 4.5 one month after receiving appropriate therapy. However, this was only assessed in 12 patients with relatively severe symptoms. The LDQ has been used in phase 2 and 3 studies of itopride.[91,92] The short-Form Leeds Dyspepsia Questionnaire (SF-LDQ)[116] was developed to be self completed and a shorter version of the original LDQ. The SF-LDQ assesses the frequency and severity over the preceding 2 months of 4 symptoms (indigestion, heartburn, regurgitation and nausea). In the validation study involving both primary and secondary care patients, Cronbach's α coefficient was 0.90, demonstrating a high level of internal consistency. Pearson's correlation coefficient for test-retest reliability 2 days apart was 0.93. Validity was demonstrated by comparison with general practitioners' diagnosis. Comparison of the summed total scores for patients in the primary care and secondary care setting yielded statistically significant differences, further confirmed the validity of the SF-LDQ. The SF-LDQ's responsiveness to change was significant in 37 patients who received treatment of known effectiveness.

C.4. Dyspepsia Symptom Severity Index (DSSI)

The DSSI[117] is a 20-item self administered questionnaire divided into 3 subscales (reflux-like, ulcer-like and dysmotility-like symptoms), quantifying the severity of dyspepsia symptoms over the past 2 weeks. The items were derived from patient focus group interviews. One global item is included at the conclusion of the questionnaire to access the patient's overall impression of dyspepsia severity. The severity questions are graded on a 5-point Likert scale ranging from 0 (absent) to 4 (very severe). Subscale internal consistency levels were high (Cronbach α 0.84-0.89) and reproducible (ICC 0.90-0.92). The DSSI correlated well with the patients' symptom diary and discriminated between patients and age-matched controls. However, despite its validity and reliability, its responsiveness was not assessed. Inclusion of GERD and nausea symptoms is another drawback. The DSSI was used in the Acotiamide U.S. Phase 2 program as a secondary endpoint.[42] Significant improvements in symptoms of bloating, postprandial nausea, and stomach pain before meals with the 300 mg dose t.i.d., were associated with significantly higher response rates on the adequate relief and OTE endpoints, and with significant improvement of several domains of the Short Form-36 quality of life scale and of the Short form Nepean Dyspepsia Index during the first weeks of the study.[42]

C.5. Other uni-dimensional dyspepsia symptom questionnaires.

The only dyspepsia questionnaire to be validated in a Chinese population, the Hong Kong Dyspepsia Index[118] originally contained 24 gastrointestinal

symptoms that were self administered, but only the 12 most discriminating symptoms (epigastric pain, upper abdominal bloating, upper abdominal dull ache, epigastric pain before meals, epigastric pain when anxious, vomiting, nausea, belching, acid regurgitation, heartburn, feeling of acidity in the stomach and loss of appetite) were selected. Symptoms are graded on a 5 point Likert scale, ranging from 1(no symptoms) to 5 (incapacitating symptoms resulting in an inability to perform daily activities and/or requiring days off work). A cut-off score of ≥16 was shown to have a sensitivity of 0.82 and specificity of 0.83 for identification of dyspeptic symptoms. There was good test-retest reliability and internal consistency, with an intraclass correlation coefficient of 0.89 and Cronbach's alpha coefficient of 0.90. Dyspepsia scores correlated negatively with all aspects of the SF-36 quality of life scale with the exception of physical functioning, lending support for its construct validity. It was also able to discriminate between patients who reported a subjective improvement in symptoms and those who reported no change or worsening. Although it assesses disease severity, this questionnaire does not measure frequency of dyspeptic symptoms, excludes some key meal-related symptoms (e.g. early satiety) and includes a number of GERD symptoms. This scoring system has been used in a study evaluating the efficacy of 4 weeks of lansoprazole treatment in patients who fulfilled Rome II criteria for FD with a baseline score ≥16.[119]

The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ)[120] is an investigator administered unidimensional 11-item instrument which evaluates 5 symptoms of NUD as defined by the Rome I consensus. The symptoms of

upper abdominal pain, nausea and upper abdominal bloating are accessed for intensity, duration and frequency; whilst vomiting and early satiety are accessed for frequency. Internal consistency of this questionnaire, as measured by Cronbach's alpha was 0.82. It was shown to be reproducible when submitted to the test-retest procedure, both by the same interviewer (ICC 0.86) and by different interviewers (ICC 0.87). Its responsiveness was demonstrated by a statistically significant reduction in the mean score obtained after treatment compared with baseline. In addition, the PADYQ demonstrated content, discriminant and criterion validity. However, the questionnaire focuses on the pain spectrum of FD, and does not address symptoms such as postprandial fullness and epigastric burning.

D. Disease specific instruments (multidimensional)

D.1. Nepean Dyspepsia Index (NDI)

The NDI[121,122] was developed primarily as a disease specific quality of life measure. From an initial 42-items to measure the impact of dyspepsia on a patient's overall quality of life, it was shortened to 25-items centred around 4 subscales: (i) interference with activities of daily living (13 items); (ii) lack of control over the illness (7 items); (iii) disturbance of eating or drinking (3 items) and (iv) sleep disturbance (2 items). A separate symptom checklist measures the frequency, severity and bothersomeness of 15 common upper gastrointestinal symptoms over a 2-week period. An expert panel documented high internal consistency for the quality of life part of the NDI (Cronbach α 0.81-0.96) as well as good discriminative and convergent validity. The NDI was used as a secondary outcome measure in phase 2 and 3 itopride studies, in tegaserod phase 2 and 3 studies and in the Acotiamide U.S. Phase 2.[42,43,91,92,123] Test retest reliability was evaluated in a multi-center study in the United States.[124] The NDI has been translated and validated in several languages, including Arabic.[125] The NDI was further shortened to 10 quality-of-life items in 5 domains[126]. This short form NDI was validated in a large multicentre trial in Europe and the United States. The short form NDI demonstrated adequate internal consistency (Cronbach α 0.71-0.76), correlated with the original long version and was highly responsive. To date,

one study[127] employed both the short and long versions of the NDI, but only the results of the long version were presented.

D.2. Severity of Dyspepsia Assessment (SODA)

The SODA[128] is a multidimensional outcome measure for dyspepsia related health that consists of 17 questions incorporating 3 dimensions of pain intensity (6 items); 7 non-pain symptoms (belching, heartburn, bloating, passing gas, sour taste, nausea and bad breath) and satisfaction scales (4 items). In this self administered questionnaire, patients are asked to assess the preceding 7 days. Reliability was good (Cronbach's α were 0.97, 0.90 and 0.92 for pain intensity, non-pain symptoms and satisfaction scales respectively) and mean change scores discriminated between patients who reported improvement compared with those who were unchanged. The psychometric properties of SODA were further evaluated in a large study of arthritis patients receiving NSAIDS or celecoxib, confirming good internal consistency.[129]

Responsiveness as evaluated by receiver operating characteristic (ROC) curves was high for pain intensity scale (area under the curve (AUC)=0.78); non-pain symptoms (AUC =0.74) and satisfaction scales (AUC=0.75). However, reproducibility was low, with ICC values of 0.49, 0.61 and 0.45 for pain intensity, non-pain symptoms and satisfaction scales respectively. Notably, SODA assesses the severity but not the frequency of symptoms, and it includes a broad range of symptoms that are also suggestive for GERD or

IBS. The SODA was used as a secondary outcome measure in the ENTER trial.[97]

D.3. Patient assessment of upper gastrointestinal symptom severity index and quality of life (PAGI-SYM and PAGI-QOL) Questionnaires.

The PAGI-SYM© questionnaire has been developed and validated for the evaluation of symptom severity and treatment responsiveness in upper gastrointestinal disorders such as FD, gastroparesis and GERD.[130] It is composed of 20 items in 6 subscales: heartburn/regurgitation (7 items), nausea/vomiting (3 items), postprandial fullness/early satiety (4 items), bloating (2 items), upper abdominal pain (2 items), and lower abdominal pain (2 items). Each item is scored from 0 (none) to 5 (very severe). The items were derived from patient focus groups and expert opinion. The PAGI-SYM questionnaire has good internal consistency, reliability and reproducibility over 2 weeks, and adequate content and construct validity in samples of subjects with upper gastrointestinal disorders.[130] Responsiveness has been evaluated in an 8-week study in GERD and FD, using OTE as a comparator.[131] The questionnaire has also been translated and validated in many different cultural and linguistic settings. Disadvantages of the PAGI-SYM questionnaire are the equal weighting of different symptoms when making up total or subscales, and the inclusion of a broad range of symptoms that are not specific for FD. However, subscales were proposed for use in specific disorders.[131] The PAGI-SYM has been used in a study evaluating the 5-HT1A agonist R137696 in FD, where no efficacy was demonstrated.[132]

The PAGI-QOL questionnaire consists of 30 items in five subscales: daily activities (10 items), clothing (2 items), diet and food (7 items), relationships (3 items) and psychological distress (8 items). Patients accord a score to each item from 0 (absent) to 5 (severely impaired). It has been validated for evaluating quality of life in FD, gastro-oesophageal reflux disease and gastroparesis by comparing it with the OTE.[133]

D.4. Glasgow Dyspepsia Severity Score (GDSS)

The GDSS[134] provides a tool for the global evaluation of dyspepsia during the preceding six months, and was shown to be reliable, valid and responsive. It records frequency of symptoms (maximum score,5); effect on daily activities (2); the number of days off work because of dyspepsia (2), frequency of medical consultations (2); home visits by a general practitioner (2); clinical investigations performed (2); use of over the counter medications (2) and prescription medications(3). Scores range from 0-20, and are significantly lower in healthy volunteers (mean score 1.16) compared to patients with duodenal ulcer (11.1) or FD (10.5). The coefficient of variation was 2% and 8% for intra- and inter-observer assessment respectively. Following eradication of *Helicobacter pylori* in duodenal ulcer, the score changed from 11.4 to 1.3, compared with an average change of 10.5 to 8.5 in patients with persistent *Helicobacter pylori* infection, confirming responsiveness of the GDSS. A score of 0 or 1 on the GDSS corresponds to complete resolution of

dyspepsia. Limitations of the GDSS are administration by an investigator, and lack of standardized definition of dyspepsia for the investigator to adhere to.

The GDSS was utilised by McColl et al[135] in a trial evaluating the efficacy of *Helicobacter pylori* eradication therapy in non ulcer dyspepsia. A Spanish translation of the GDSS and a Likert-scale symptomatic test were evaluated for responsiveness to treatment, validity and reproducibility by means of phone interview.[136] Both the GDSS and Likert-scales were valid (higher scores in patients undergoing endoscopy than in healthy controls); had adequate response after subjects completed *Helicobacter pylori* eradication therapy, and showed low intraobserver variation whether the interviews were conducted by phone or by a combination of phone and clinical interview. The GDSS was superior to the Likert-scale in remaining reproducible even when conducted by different observers.

D.5. Other disease specific multidimensional instruments.

The Clinical Dyspepsia Questionnaire is a condition specific self-administered measure of dyspepsia and peptic-ulcer related symptoms. It assesses the frequency and severity of symptoms and their impact on QOL.[137] Patient scores on this questionnaire reviewed only small to moderate correlations (0.20 to 0.66) with the SF-36 quality of life survey, but further validity tests showed significant correlations with general practitioner perceptions of symptom severity, family history of peptic ulcer disease and whether patients

were referred. Test-retest reliability yielded an intraclass coefficient value of 0.69. However, responsiveness was not assessed.

The Gastrointestinal Symptom Score (GIS) is a 10-item questionnaire developed for the study of FD.[138] It was developed based on patient focus groups, and validated by comparison with the Nepean Dyspepsia Index. However, the questionnaire includes typical GERD symptoms like heartburn, regurgitation and retrosternal discomfort. The GIS has been used in studies of herbal medicines in FD.[139-141]

Quality of life measures

Quality of life (QOL) measures can be assessed either by generic instruments or by disease specific measures. The Short Form of General Health Questionnaire (SF-36)[142] is a generic instrument that has been widely used in various disorders. To evaluate more specifically QOL in specific disorders, such as in functional gastrointestinal disorders, questionnaires that are available include the Nepean Dyspepsia Index (NDI)[121-126] for FD, the PAGI-QOL for upper GI disorders[130] and the IBS-QOL[143] for IBS. Although these quality of life measures have been strongly recommended as secondary outcome measures, they are not utilised as primary outcome measures as they are regarded by regulatory agencies as being insufficiently responsive to treatment, although this may not necessarily be correct.[10]

Various trials have measured QOL as primary or secondary outcome measures (see Table 1). In a recent trial of tegaserod for dysmotility-type FD, significant improvement in treatment satisfaction and reduced work and daily activity impairment were reported in the treatment group compared with placebo. [144]

DISCUSSION

Evaluating the therapeutic efficacy of drugs in FD remains challenging due to the variable intensity of symptoms in one individual and the high placebo response rates.[13,145-149] The choice of the endpoint for a clinical trial is a key factor in assessing therapeutic efficacy. However, there is a lack of consensus on the best method to measure clinical outcome in FD drug trials.[13,14] In this setting, the most recent mandate emphasises the use of PRO in symptom reporting.[9], The dichotomous endpoint of adequate or satisfactory relief has seen increasing use as primary endpoint in recent trials, while alternative outcome measures such as integrative questionnaires are also considered potentially acceptable, but these views seem largely driven by experience with IBS trials.[26-37]

Achieving the adequate relief endpoint in clinical drug trials was instrumental in obtaining FDA approval initially for ranitidine for the treatment of heartburn,[151] for alosetron in IBS[26-28], and subsequently for tegaserod in IBS (using satisfactory relief).[34-37] Evidence supporting responsiveness of the adequate or satisfactory relief endpoints was documented mainly in IBS trials, where the binary endpoint correlated with other standard measures of treatment efficacy.[26-37,152,153]

On the other hand, limitations of the binary endpoints include the lack of sensitivity and validation, inability to detect worsening of symptoms (unlike the OTE), lack of information on the magnitude of improvement needed to reach

adequate relief, the potential for failure to report adequate relief in subsequent weeks once achieved, the influence of baseline severity of symptoms and response outcomes and variable interpretation of the adequate relief construct.[39] Based on an observational IBS study in a health maintenance organization setting, it has been suggested that patients with mild symptoms at baseline are more likely to achieve satisfactory relief compared to those with severe symptoms at baseline.[154] Contrary to this suggestion, the greatest improvement in percentage of days that FD patients reported satisfactory relief of symptoms occurred in those who had severe baseline symptoms compared with those who had mild symptoms ($p<0.05$).[43] Hence, the use of binary relief endpoints in FD deserves additional validation studies. Compared to questionnaires that use specific FD symptoms, the binary or overall treatment evaluation question may be more suitable to capture an overall treatment impact in functional disorders, especially as this approach has the potential to also contain changes in associated more global symptoms (e.g. fatigue).

In addressing the use of an acceptable alternative outcome measure, a questionnaire should undergo appropriate psychometric validation in accordance with the FDA guidance regulations and Rome III guidelines on clinical trials.[9,10] In our opinion, existing questionnaires present some methodological flaws, mainly due to the failure to satisfy the psychometric property of content validity. In the light of the Rome III criteria[4], none of the existing questionnaires are sufficiently validated to be recommended

unequivocally as the primary outcome measure for FD trials. Central to the problem of previously validated questionnaires was the inclusion of many symptoms which are not considered cardinal FD symptoms according to current criteria[4]. An adequate and valid questionnaire, i.e. one that incorporates all the psychometric validation standards, in particular incorporating the Rome III criteria for FD, would facilitate future therapeutic trials in FD.

The absence of a validated instrument for the evaluation of treatment efficacy in FD according to Rome III criteria and the lack of robust psychometric validation data for the binary outcome measure in FD trials should not impair ongoing or planned drug development in this area. One plausible solution is to build upon the heritage of current instruments to serve as a useful framework in designing a new questionnaire. Existing questionnaires that include many non-FD symptoms can be narrowed down to evaluate only the cardinal symptoms according to Rome III. Such an approach was used in the itopride phase 2 and 3 studies, which used 2 FD-specific questions (severity of epigastric pain and of postprandial fullness) from the LDQ as co-primary endpoint, whereas the full LDQ score was a secondary endpoint.[91,92] The most valuable questionnaires are probably the ones that used patient focus groups to identify relevant symptoms, such as the DSSI, PAGI-SYM or GIS questionnaires.[117,121,130,138] Another possibility is to incorporate the OTE with the binary endpoints of adequate or satisfactory relief, for which there is extensive experience from IBS trials[26-37] and emerging experience from FD trials.[25,40-43] The OTE was extensively used in the OCAY, CADET and ENTER

trials[89,94-97] and most recently in the tegaserod and acotiamide studies.[42,43] The subjective nature of the global outcome approach or OTE allows the individual to integrate all aspects of his condition into a single treatment outcome. Although the OTE is particularly suitable to show deterioration, it is prone to recall bias as it compares current symptoms to pre-treatment severity. Nevertheless, using both the OTE and binary outcome to evaluate treatment efficacy would provide valuable information both from a global perspective as well as in assessing symptom improvement and/or deterioration.

Similar to other FGIDs, the magnitude of the placebo response is a major concern in the evaluation of efficacy of drugs in FD. The available recent studies do not show a major difference in responder rate when comparing binary endpoints to OTE responses or to responder rates on composite questionnaires like the LDQ or DSSI, provided the cut-off level for response was adequately chosen.[42,43,92] Indeed, these studies illustrate how the magnitude of the placebo effect depends on the level of improvement used to define a responder. For instance, in the itopride phase 3 trials, an improvement of 1 point on the LDQ had a placebo response of 70 to 75%; when a 2 point improvement was chosen as response definition, the placebo response was 45 to 55%, quite similar to the placebo response on a global patient assessment of efficacy.[92] Similarly, requiring a 50% weekly response rate over a 75% weekly response rate is associated with a higher placebo rate.[42,43] In addition, the tegaserod phase 3 studies showed a high placebo response rate, and a low therapeutic gain of active drug over placebo, when

patients with mild symptoms were considered, while a bigger therapeutic gain over placebo was obtained in those with moderate or severe symptoms.[43] Hence, in order to minimize the placebo effect, inclusion of patients with at least moderate symptom severity, and the choice of a high enough threshold for response definition is recommended.

Some studies in FD have used elimination of symptoms, usually measured on Likert scales, as an endpoint.[98,114,135] Although it is clear that becoming asymptomatic is clinically relevant, the proportion of patients with functional gastrointestinal disorders in whom this can be achieved by a pharmacological intervention (and by the comparator placebo arm) is likely to be low. Placebo responses are indeed low when this endpoint is considered. [98,114,135] Elimination of symptoms, therefore, probably sets a very high threshold for response and focusing on this endpoint induces a risk of disregarding clinically relevant degrees of symptom improvement in FD and other FGIDs. International consensus guidelines, as well as FDA guidance, therefore focus on identifying and recognizing significant symptom improvement using validated scales,[9-13] rather than symptom elimination which can still be considered a useful secondary outcome variable.

No regulatory guideline exists on which magnitude of an active drug response rate over placebo is needed to be considered clinically relevant. This was also not specifically addressed by the Rome III working committee on design of treatment trials in functional disorders.[10] However, a group of experts proposed a minimal range of efficacy of 10 to 15% over placebo to be

clinically relevant for functional disorders and for the irritable bowel syndrome in particular.[155] It seems reasonable to assume that the same magnitude of margin over placebo could also be considered clinically relevant in FD.

In conclusion, there are emerging data on the use of binary outcome measures in FD clinical trials, although further validation is required. Most existing questionnaires do not sufficiently fulfil psychometric validation criteria, in view of the latest consensus definitions of FD by Rome III criteria. Hence, evaluating and comparing the various options will need to go hand in hand with the analysis of recent, ongoing or future pharmaceutical trials for the treatment of FD, and optimization of potential endpoints that could be derived from existing instruments.

Grant support:

Pieter Janssen is a research fellow of the FWO Flanders. This work was supported by a Methusalem grant to Jan Tack, M.D., Ph.D.

REFERENCES

1. Management of dyspepsia: Report of a working party. Lancet1988;1:576-9.
2. Talley NJ, Koch KL, Koch M, et al. Functional dyspepsia: A classification with guidelines for diagnosis and management. Gastrointest Int 1991;4:145-60.
3. Talley NJ, StanghelliniV, Heading RC, Koch KL, Malagelada JR, Tytgat GNJ. Functional gastroduodenal disorders. Gut 1999; 45(Suppl II):37-42.
4. Tack J, Talley NJ, Camilleri M, Holtmann G, Hu PJ, Malagelada JR, Stanghellini V. Functional gastroduodenal disorders. Gastroenterology 2006;130:1466-79.
5. Drossman DA, Thompson G, Talley NJ, et al. Identification of subgroups of functional gastrointestinal disorders. Gastroenterol Int 1991;4:145-60.
6. Talley NJ, Phillips SF. Non-ulcer dyspepsia: Potential causes and pathophysiology. Ann Intern Med 1988;108:965-79.
7. Talley NJ, Weaver AL, Zinsmeister AR, et al. Onset and disappearance of gastrointestinal symptoms and functional gastrointestinal disorders. Am J Epidemiol 1992;136:165-77.
8. Shaib Y, El-Serag HB. The prevalence and risk factors of functional dyspepsia in a multiethnic population in the United States. Am J Gastroenterol 2004;99:2210-6.

9. US Food and Drug Administration (FDA). Guidance for Industry. Patient reported outcome measures: use in medical product development to support labeling claims. 2006.

10. Design of Treatment Trials Committee; Irvine EJ, Whitehead WE, Chey WD, Matsueda K, Shaw M, Talley NJ, Veldhuyzen Van Zanten SJ. Design of treatment trials for functional gastrointestinal disorders. Gastroenterology 2006;130:1538-51.

11. Veldhuyzen van Zanten SJ. A systematic overview (meta-analysis) of outcome measures in Helicobacter pylori gastritis trials and functional dyspepsia. Scan J Gastroenterol 1993;28(Suppl 199):40-43.

12. Talley NJ. A critique of therapeutic trials in Helicobacter pylori-positive functional dyspepsia. Gastroenterology 1994;106:1174-83

13. Veldhuyzen van Zanten SJ, Cleary C, Talley NJ, et al. Drug treatment of functional dyspepsia: a systematic analysis of trial methodology with recommendations for design of future trials. Am J Gastroenterol 1996;91:660-73.

14. Veldhuyzen van Zanten SJO. Assessment of outcome in dyspepsia: has progress been made? Gut 2002;50(Suppl IV):23-25

15. Tack J, Bisschops R, Sarnelli G. (2004). Pathophysiology and treatment of functional dyspepsia. Gastroenterology 127: 1239-55.

16. Guyatt GH, Veldhuyzen van Zanten SJO, Feeney DH, et al. Measuring quality of life in clinical trials: a taxonomy and review. Can Med J Assoc 1989;140:1441-8.

17. Talley NJ, Nyren O, Drossman DA, Heaton KW, Veldhuyzen van Zanten SJO, Koch MM, Ransohoff DF. The irritable bowel syndrome:

toward optimal design of controlled treatment trials. Gastroenterol Int 1993;6:189-211.

18. Guyatt GH, Townsend M, Berman LB, Keller JL. A comparison of Likert and Visual Analogue Scales for measuring change in function. Chron Dis 1987;40:1129-33.

19. Jaeschke R, Singer J, Guyatt GH. Measurement of health status:ascertaining the minimal clinically important difference. Control Clin Trials 1989;10:407-15.

20. Guyatt G, Deyo RA, Charlson M, Levine MN, Mitchell A. Responsiveness and validity in health status measurement: a clarification. J Clin Epidemiol 1989;42:403-8.

21. Veldhuyzen van Zanten SJ, Talley NJ, Bytzer P, Klein KB, Whorwell PJ, Zinsmeister AR. Design of treatment trials for functionalgastrointestinal disorders. Gut 1999; 45(Suppl 2):1169-77.

22. Hays RD, Hadorn D. Responsiveness to change: an aspect of validity, not a separate dimension. Qual Life Res 1992;1:73-5.

23. Kirshner B, Guyatt GH. A methodological framework for assessing health indices. J Chron Dis 1985;38:27-36.

24. Cook KF, Rabeneck L, Campbell CJM, Wray NP. Evaluation of a multidimensional measure of dyspepsia related health for use in a randomized clinical trial. J Clin Epidemiol 1999;52:381-92.

25. Farup PG, Hovde O, Torp R, Wetterhus S. Patients with functional dyspepsia responding to omeprazole have a characteristic gastro-oesophageal reflux pattern. Scan J Gastroenterol 1999;34:575-9.

26. Camilleri M, Mayer EA, Drossman DA, Heath A, Dukes GE, McSorley D, Kong S, Mangel AW, Northcutt AR. Improvement in pain and bowel function in female irritable bowel patients with alosetron, a 5-HT3 receptor antagonist. Aliment Pharmacol Ther 1999;13:1149-59.

27. Camilleri M, Northcutt SR, Kong S, Dukes GE, McSorley D, Mangel AW. Efficacy and safety of alosetron in women with irritable bowel syndrome: a randomized, placebo-controlled trial. Lancet 2000;355:1035-40.

28. Camilleri M, Chey WY, Mayer EA, Northcutt AR, Heath A, Dukes GE, McSorley D, Mangel AM. A randomized controlled clinical trial of the serotonin type 3 receptor antagonist alosetron in women with diarrhoea-predominant irritable bowel syndrome. Arch Intern Med 2001;161:1733-40.

29. Miner P, Stanton DB, Carter F, et al. Cilansetron in irritable bowel syndrome with diarrhoea predominance (IBS-D); efficacy and safety in a 3 month US study. Am J Gastroenterol 2004;99:S277.

30. Leventer SM, Kucharik RF, Keogh JC, et al. The potential dextofisopam for treatment of irritable bowel syndrome and the irritable bowel syndrome. Am J Gastroenterol 2004;99:S279.

31. Bradette M, Moennikes H, Carter F, Krause G, Caras S, Steinborn C. Cilansetron in irritable bowel syndrome with diarrhoea predominance (IBS-D): efficacy and safety in a 6 month global study. Gastroenterology 2004;126(Suppl. 2):A42.

32. Coremans G, Clouse RE, Carter F, et al. Cilansetron, a novel 5-HT3 antagonist demonstrated efficacy in males with irritable bowel

syndrome with diarrhoea-predominance (IBS-D). Gastroenterology 2004;126(4 Suppl.2):A643.

33. Leventer S, Raudibaugh K, Frissora C, et al. The safety and efficacy of dextofisopam in patients with diarrhoea-predominant or alternating irritable bowel syndrome. Gastroenterology 2005;128(4 Suppl 2):A94.

34. Kellow J, Lee OY, Chang FY, et al. An Asia-Pacific, double-blind, placebo controlled, randomized study to evaluate the efficacy, safety, and tolerability of tegaserod on patients with irritable bowel syndrome. Gut 2003;52:671-6.

35. Nyhlin H, Bang C, Elsborg L, Silvennoinen J, Holme I, Ruegg P, Jones J, Wagner A. A double-blind, placebo-controlled, randomized study to evaluate the efficacy, safety and tolerability of tegaserod in patients with irritable bowel syndrome. Scand J Gastroenterol 2004;39:119-126.

36. Muller-Lissner S, Holtmann G, Rueegg P, Weidinger G, Loffler H. Tegaserod is effective in the initial and re-treatment of irritable bowel syndrome with constipation. Aliment Pharmacol Ther 2005;21:11-20.

37. Tack J, Muller-Lissner S, Bytzer P et al. A randomized controlled trial assessing the efficacy and safety of repeated tegaserod therapy in women with irritable bowel syndrome with constipation. Gut 2005; 54:1707-13.

38. Mangel AW. Personal view: adequate relief as a primary endpoint in irritable bowel syndrome. Aliment Pharm Ther 2006; 23:879-881.

39. Camilleri M, Drossman DA, Fehnel S, Mangel A, Mayer EA, Talley NJ. Primary endpoints for FGID trials: A review of binary and integrated symptom assessments. Clin Gastro Hepatol 2007;5:534-40.

40. Talley NJ, Van Zanten SV, Saez LR, Dukes G, Perschy T, Heath M, Kleoudis C, Mangel AW. A dose ranging placebo controlled randomized trial of alosetron in patients with functional dyspepsia. Aliment Pharmacol Ther 2001;15:525-37.

41. Matsueda K, Hongo M, Sasaki D, Kusano M, Harasawa S, Arakawa T, Haruma K, Nakashima M, Miwa T and Saitou Y. Therapeutic Efficacy of Novel Agent(Z-338) in Functional Dyspepsia(FD). Gastroenterology 2006, A467 (abstract).

42. Talley NJ, Tack J., Kowalski D., Borton M.A., and Barve A. A Novel Acetylcholine Esterase Inhibitor Acotiamide Hydrochloride (YM443) in Functional Dyspepsia: Efficacy in a Randomized, Double-Blind, Placebo-Controlled Dose Ranging Trial. Gastroenterology 2008; 4, suppl. 1: A-157 (abstract).

43. Vakil N, Talley NJ, Zakko S, Laine L, Chey W, Tack J. Tegaserod treatment for functional dyspepsia dysmotility symptoms: results of two randomized, controlled trials. Am J Gastroenterol. 2008 Aug;103(8):1906-19.

44. Jian R, Ducrot F, Ruskone A, et al. Symptomatic, radionuclide and therapeutic assessment of chronic idiopathic dyspepsia. A double-blind placebo-controlled evaluation of cisapride. Dig Dis Sci 1989;34:657-64.

45. Tatsuta M, Iishi H, Nakaizumi A, Okuda S. Effect of treatment with cisapride alone or in combination with domperidone on gastric emptying and gastrointestinal symptoms in dyspeptic patients. Aliment Pharmacol Ther 1992; 6:221-8.

46. Sarin SK, Sharma P, Chawla YK, et al. Clinical trial on the effect of domperidone on non-ulcer dyspepsia. Indian J Med Res 1986; 83:623-8.

47. Champion MC, MacCannell KI, Thomson AB, et al. A double-blind randomized study of cisapride in the treatment of non-ulcer dyspepsia. The Canadian Cisapride Nud Study Group. Can J Gastroenterol 1997;11:127-34.

48. De Groot GH, De Both PSM. Cisapride in functional dyspepsia in general practice. A placebo-controlled, randomized, double-blind study. Aliment Pharmacol Ther 1997;11:193-9.

49. Tack J, Masclee A, Heading R, et al. A dose-ranging, placebo-controlled, pilot trial of acotiamide in patients with functional dyspepsia. Neurogastroenterol Motil. 2009 Mar;21(3):272-80.

50. Gotthard R, Bodemar G, Brodin U, Jonsson KA. Treatment with cimetidine, antacid or placebo in patients with dyspepsia of unknown origin. Scand J Gastroenterol 1988;23:7-18.

51. Bekhti A, Rutgeerts L. Domperidone in the treatment of functional dyspepsia in patients with delayed gastric emptying. Postgrad Med J1979;55(Suppl.1):30-32.

52. Van de Mierop L, Rutgeerts L, Van den Langenbergh B, Staessen A. Oral domperidone in chronic postprandial dyspepsia. Digestion 1979;19:244-50.

53. Haarmann K, Lebkuchner F, Widmann A, Kief W, Esslinger M. A double-blind study of domperidone in the symptomatic treatment of chronic post-prandial upper gastrointestinal distress. Postgrad Med J 1979;55(Suppl. 1):24-7.

54. Lance P, Wastell C, Schiller KFR. A controlled trial of cimetidine for the treatment of non-ulcer dyspepsia. J Clin Gastroenterol 1986;8:414-8.

55. Smith PM, Troughton AH, Gleeson F, Walters J, McCarthy CF. Pirenzepine in non-ulcer dyspepsia: a double-blind multicentre trial. J Int Med Res 1990;18:16-20.

56. Deruyttere M, Lepoutre L, Heylen H, Samain H, Pennoit H. Cisapride in the management of chronic functional dyspepsia: A multicentre, double-blind, placebo-controlled study. Clin Ther 1987;10:44-51.

57. Rosch W. Cisapride in non ulcer dyspepsia. Results of a placebo-controlled trial. Scand J Gastroenterol 1987;22:161-4.

58. Hannon R. Efficacy of cisapride in patients with nonulcer dyspepsia. A placebo-controlled study. Curr Ther Res Clin Exp 1987;42:814-22.

59. Goethals C, Van De Mierop L. Cisapride in the treatment of chronic functional dyspepsia. Results of a double-blind, placebo-controlled, crossover study. Curr Ther Res 1987;42:261-7.

60. Francois I, De Nutte N. Nonulcer dyspepsia: effect of the gastrointestinal prokinetic drug cisapride. Curr Ther Res 1987;41:891-8.

61. De Nutte N, Van Ganse W, Witterhulghe M, Defrance P. Relief of epigastric pain in nonulcer dyspepsia: controlled trial of the promotility drug cisapride.Clin Ther 1989;11:62-8.

62. Wood SF, Penney SC, Cochran KM. Cisapride in functional dyspepsia: A double-blind placebo-controlled randomized trial in general practice patients. Scand J Gastroenterol 1993;28(Suppl 195):5-10.

63. Chung JM. Cisapride in chronic dyspepsia: results of a double-blind, placebo-controlled trial. Scand J Gastroenterol 1993;195(Suppl)11-4.

64. Van Outryve M, De Nutte N, Van Eeghem P, Gooris JP. Efficacy of cisapride in functional dyspepsia resistant to domperidone or metoclopramide: A double-blind placebo-controlled study. Scand J Gastroenterol 1993;28(Suppl. 195):47-53.

65. Fumagalli I, Hammer B. Cisapride versus metoclopramide in the treatment of functional dyspepsia. A double-blind comparative trial. Scand J Gastroenterol 1994;29:33-7.

66. Al-Quorain A, Larbi EB, Al-Shedoki F. A double-blind, randomized, placebo-controlled trial of cisapride in Saudi Arabs with functional dyspepsia. Scand J Gastroenterol 1995;30:531-4.

67. Kellow JE, Cowan H, Shuter B, et al. Efficacy of cisapride therapy in functional dyspepsia. Aliment Pharmacol Ther 1995;9:153-60.

68. Carvalhinhos A, Fidalgo P, Freire A, Matos L. Cisapride compared with ranitidine in the treatment of functional dyspepsia. Euro J Gastroenterol and Hepatol 1995;7:411-7.

69. Yeoh KG, Kang JY, Tay HH, et al. Effect of cisapride on functional dyspepsia in patients with and without histological gastritis: a double-blind placebo controlled trial. J Gastroenterol Hepatol 1997;12:13-8.

70. Hansen JM, Bytzer P, Schaffalitzky de Muckadell O. Placebo-controlled trial of cisapride and nizatidine in unselected patients with functional dyspepsia. Am J Gastroenterol 1998;93:368-74.

71. Creytens G. Effect of the non-antidopaminergic drug cisapride on postprandial nausea. Curr Ther Res 1999;36(5):1063-70.

72. Chen JD, Ke MY, Lin XM, Wang Z, Zhang M. Cisapride provides symptomatic relief in functional dyspepsia associated with gastric myoelectrical abnormality. Aliment Pharmacol Ther 2000;14:1041-7.

73. Fraitag B, Homerin M, Hecketsweiler P. Double-blind dose-response multicenter comparison of fedotozine and placebo in treatment of nonulcer dyspepsia. Dig Dis Sci 1994;39:1072-77.

74. Read NW, Abitol JI, Bardhan KD, Whorwell PJ, Fraitag B. Efficacy and safety of the peripheral kappa agonist fedotozine versus placebo in the treatment of functional dyspepsia. Gut 1997;41:664-8.

75. Delattre M, Malesky M, Prinzie A. Symptomatic treatment of non-ulcer dyspepsia with cimetidine. Curr Ther Res 1985;37(5):980-1.

76. Kelbaek H, Linde J, Eriksen J, et al. Controlled clinical trial of treatment with cimetidine for non-ulcer dyspepsia. Acta Med Scand 1985;217:281-7.

77. Singal AK, Kumar A, Broor SL. Cimetidine in the treatment of non-ulcer dyspepsia: results of a randomized double-blind, placebo-controlled study. Curr Med Res Opin 1989;11:390-7.

78. Kairaluoma MI, Hentilae R, Alavaikko M, et al. Sucralfate versus placebo in treatment of non-ulcer dyspepsia. Am J Med 1987; 83:51-5.

79. Loffeld RJ, Potters HV, Stobberingh E, Flendrig JA, van Spreeuwel JP, Arends JW. Campylobacter associated gastritis in patients with non-ulcer dyspepsia: a double blind placebo controlled trial with colloidal bismuth subcitrate. Gut 1989;30:1206-12.

80. Vaira D, Holton J, Ainley C, Falzon M, Osborn J, D'Anna L, et al. Double blind trial of colloidal bismuth subcitrate versus placebo in Helicobacter pylori positive patients with non-ulcer dyspepsia. Italian J Gastroenterol 1992;24(7):400-4.

81. Hausken T, Stene-Larsen G, Lange O, et al. Misoprostol treatment exacerbates abdominal discomfort in patients with non-ulcer dyspepsia and erosive pyloric changes. A double-blind, placebo-controlled multicentre study. Scand J Gastroenterol 1990;25:1028-33.

82. Bolling-Sternevald E, Lauritsen K, Aalykke C, et al. Effect of profound acid suppression in functional dyspepsia: a double-blind, randomized, placebo-controlled trial. Scand J Gastroenterol 2002;37:1395-1402.

83. Peura DA, Kovacs TO, Metz D, Siepman N, Pilmer BL, Talley NJ. Lansoprazole in the treatment of functional dyspepsia: Two double-blind, randomized, placebo-controlled trials. Am J Med 2004;116:740-8.

84. Blum AL, Arnold R, Stolte M, Fischer M, Koelz HR and the Frosch Study Group. Short course acid suppressive treatment for patients

with functional dyspepsia: results depend on Helicobacter pylori status. Gut 2000;47:473-80.

85. Talley NJ, Verlinden M, Snape W, Beker JA, et al. Failure of a motilin receptor agonist (ABT-229) to relieve the symptoms of functional dyspepsia in patients with and without delayed gastric emptying: a randomized double-blind placebo-controlled trial. Aliment Pharmacol Ther 2000;14:1653-61.

86. Hallerback BI, Bommelaer G, Bredberg E, et al. Dose finding study of mosapride in functional dyspepsia: a placebo-controlled, randomized study. Aliment Pharmacol Ther 2002;16:959-67.

87. Holtmann G, Adam B, Haag S, Collet W, Grunewald E, Windeck T. Efficacy of artichoke leaf extract in the treatment of patients with functional dyspepsia: a six-week placebo-controlled, double-blind multicentre trial. Aliment Pharmacol Ther 2003; 18:1099-1105.

88. Talley NJ, Meineche-Schmidt V, Pare P, et al. Efficacy of omeprazole in functional dyspepsia: double-blind, randomized, placebo-controlled trials (the Bond and Opera studies). Aliment Pharmacol Ther1998;12:1055-65.

89. Blum, AL, Talley NJ, O'Morain C, et al. Lack of effect of treating Helicobacter pylori infection in patients with nonulcer dyspepsia. Omeprazole plus Clarithromycin and Amoxicillin Effect One Year after Treatment (OCAY) Study Group. N Engl J Med 1998;339:1875-81.

90. Talley NJ, Janssens J, Lauritsen K, et al. Eradication of Helicobacter pylori in functional dyspepsia: Randomised double blind placebo controlled trial with 12 months' follow up. The Optimal Regimen Cures

Helicobacter Induced Dyspepsia (ORCHID) Study group. BMJ 1999;318:833-7.

91. Holtmann G, Talley NJ, Liebregts T, Adam B, Parow C. A placebo-controlled trial of itopride in functional dyspepsia. N Eng J Med 2006;354:832-40.

92. Talley NJ, Tack J, Ptak T, Gupta R, Giguere M. Itopride in Functional Dyspepsia: Results of two phase III multicenter, randomized, double-blind, placebo-controlled trials. Gut 2008.

93. Veldhuyzen van Zanten SJO, Tytgat KMAJ, Pollak PT, Goldie J, Goodacre RL, Riddell RH, Hunt RH. Can severity of symptoms be used as outcome measures in trials of non-ulcer dyspepsia and Helicobacter pylori associated gastritis? J Clin Epidemiol 1993;46:273-9.

94. Veldhuyzen van Zanten SJO, Chiba N, Armstrong D, et al. A randomized trial comparing omeprazole, ranitidine, cisapride, or placebo in Helicobacter pylori negative, primary care patients with dyspepsia:The CADET-HN study. Am J Gastroenterol 2005;100:1477-88.

95. Chiba N, Veldhuyzen van Zanten SJO, Sinclair P, et al. Treating Helicobacter pylori infection in primary care patients with uninvestigated dyspepsia: The Canadian adult dyspepsia empiric treatment-Helicobacter pylori positive (CADET-HP) randomised controlled trial. BMJ 2002;324:1012-6.

96. Armstrong D, Veldhuyzen van Zanten SJO, Barkun A, et al, the CADET-HR Study Group. Heartburn-dominant, uninvestigated

dyspepsia: A comparison of 'PPI-Start' and 'H2-RA-Start' management strategies in primary care. The CADET-HR Study. Aliment Pharmacol Ther 2005;21:1189-202.

97. Veldhuyzen van Zanten S, Armstrong D, Chiba N, et al. Esomeprazole 40 mg once a day in patients with functional dyspepsia: The Randomized, Placebo-Controlled "ENTER" Trial. Am J Gastroenterol

98. Matsueda K., Hongo M., Sasaki D., Kusano M., Harasawa S;, Arakawa T., Haruma K., Miwa T., and Nkashima M. Clinical trial: dose-dependent therapeutic efficacy of acotiamide hydrochloride (Z-338) in patients with functional dyspepsia - 100 mg t.i.d. is an optimal . Neurogastroenterol Motil. 2010 Jun;22(6):618-e173.

99. Weberg R, Berstad A. Low dose antacids and pirenzipine in the treatment of patients with nonulcer dyspepsia and erosive prepyloric changes. A randomized, double-blind, placebo-controlled trial. Scand J Gastroenterol 1988;23:237-43.

100. Johannessen T, Fjosne U, Kleveland PM, et al. Cimetidine responders in non-ulcer dyspepsia. Scan J Gastroenterol 1988;23:327-336.

101. Seno H, Nakase H, Chiba T. Usefulness of famotidine in functional dyspepsia patient treatment: comparison among prokinetic, acid suppression and antianxiety therapies. Aliment Pharmacol Ther 2005;21(Suppl2):32-36.

102. Hausken T, Berstad A. Cisapride treatment of patients with non-ulcer dyspepsia and erosive prepyloric changes. A double-blind, placebo-controlled trial. Scan J Gastroenterol 1992;27:213-7.

103. May B, Kohler S, Schneider B. Efficacy and tolerability of a fixed combination of peppermint oil and caraway oil in patients suffering from functional dyspepsia. Aliment Pharmacol Ther 2000;14:1671-77.

104. Holtmann G, Gschossmann J, Mayr P, Talley NJ. A randomized placebo-controlled trial of simethicone and cisapride for the treatment of patients with functional dyspepsia. Aliment Pharmacol Ther 2002;16:1641-8.

105. Veldhuyzen van Zanten SJO, Chiba N, Armstrong D, et al. Validation of a 7-point global overall symptom (GOS) scale to measure the severity of dyspepsia symptoms in clinical trials. Aliment Pharmacol Ther 2006;23:521-9.

106. Armstrong D, Veldhuyzen van Zanten SJO, Barkun AN, et al. Symptom response at 1-week predicts responders at 4-weeks with esomeprazole (40 mg bid and od) in patients with uninvestigated dyspepsia. Am J Gastroenterol 2003;98(9 Suppl1):S48-9.

107. Wiklund IK, Junghard O, Grace E, et al. Quality of life in reflux and dyspepsia patients. Psychometric documentation of a new disease-specific questionnaire (QOLRAD). Eur J Surg 1998;583:41-9.

108. Svedlund J, Sjodin I, Dotevall G. GSRS-a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 1988;33:129-34.

109. Shaw MJ, Talley NJ, Beebe TJ, et al. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. Am J Gastroenterol 2001;96:52-7.

110. Dimenas E, Glise H, Hallerback B, et al. Quality of life in patients with upper gastrointestinal symptoms. An improved evaluation of treatment regimens? Scand J Gastroenterol 1993;28:681-7.

111. Revicki DA, Wood M, Wiklund I, et al. Reliability and validity of the Gastrointestinal Symptom Rating Scale in patients with gastroesophageal reflux disease. Qual Life Res 1998;7:75-83.

112. Talley NJ, Riff DS, Schwartz H, Marcuard SP. Double-blind placebo-controlled multicentre studies of rebamipide, a gastroprotective drug, in the treatment of functional dyspepsia with or without Helicobacter pylori infection. Aliment Pharmacol Ther 2001;15:1603-11.

113. Bolling-Sternevald E, Lauritsen K, Aalykke C, et al. Effect of profound acid suppression in functional dyspepsia: a double-blind, randomized, placebo-controlled trial. Scand J Gastroenterol 2002;37:1395-1402.

114. Miwa H, Nagahara A, Tominaga K, Yokoyama T, Sawada Y, Inoue K, Ashida K, Fukuchi T, Hojo M, Yamashita H, Tomita T, Hori K, Oshima T. Efficacy of the 5-HT1A agonist tandospirone citrate in improving symptoms of patients with functional dyspepsia: a randomized controlled trial. Am J Gastroenterol. 2009 Nov;104(11):2779-87.

115. Moayyedi P, Duffett S, Braunholtz D, Mason S, Richards ID, Dowell AC, Axon AT. The Leeds Dyspepsia Questionnaire:a valid tool for measuring the presence and severity of dyspepsia. Aliment Pharmacol Ther 1998;12:1257-62.

116. Fraser A, Delaney BC, Ford AC, Qume M, Moayyedi P. The Short-Form Leeds Dyspepsia Questionnaire validation study. Aliment Pharmacol Ther 2007;25:477-86.

117. Leidy NK, Farup C, Rentz AM, et al. Patient-based assessment in dyspepsia: Development and validation of Dyspepsia Symptom Severity Index (DSSI). Dig Dis Sci 2000; 45:1172-9.

118. Hu WM, Lam KF, Wong YH, et al. The Hong Kong index of dyspepsia: A validated symptom severity questionnaire for patients with dyspepsia. J Gastroenterol Hepatol 2002;17:545-51.

119. Wong WM, Wong BCY, Hung WK, et al. Double blind, randomised placebo controlled study of four weeks of lansoprazole for the treatment of functional dyspepsia in Chinese patients. Gut 2002;51:502-6

120. Sander GB, Mazzoleni LE, Francesconi CF, et al. Development and validation of a cross-cultural questionnaire to evaluate nonulcer dyspepsia: The Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ). Dig Dis Sci 2004;49:1822-9.

121. Talley NJ, Haque M, Wyeth JW, Stace NH, Tytgat GNJ, Stanghellini V, Holtmann G, Verlinden M, Jones M. Development of a new dyspepsia impact scale: the Nepean Dyspepsia Index. Aliment Pharmacol Ther 1999;13:225-35.

122. Talley NJ, Verlinden M, Jones MP. Validity of a new quality of life scale for functional dyspepsia: a United States multicentre trial of the Nepean Dyspepsia Index (NDI). Am J Gastroenterol 1999;94: 2390-7.

123. Functional dyspepsia, delayed gastric emptying, and impaired quality of life. Gut. 2006 Jul;55(7):933-9

124. Validity of a new quality of life scale for functional dyspepsia: a United States multicenter trial of the Nepean Dyspepsia Index. Am J Gastroenterol. 1999 Sep;94(9):2390-7

125. Initial linguistic and psychometric validation of the Arabic version of Nepean Dyspepsia Index. Saudi Med J. 2006 Oct;27(10):1554-60.

126. Talley NJ, Verlinden M, Jones M. Quality of life in functional dyspepsia: responsiveness of the Nepean Dyspepsia Index and development of a new 10-item short form. Aliment Pharmacol Ther 2001;15:207-16.

127. Holtmann G, Adam B, Haag S, Collet W, Grunewald E, Windeck T. Efficacy of artichoke leaf extract in the treatment of patients with functional dyspepsia: a six-week placebo-controlled, double-blind multicentre trial. Aliment Pharmacol Ther 2003; 18:1099-1105.

128. Rabeneck L, Cook KR, Wristers K, Souchek J, Menke T, Wray NP. SODA (Severity of Dyspepsia Assessment): A new effective outcome measure for dyspepsia-related health. J Clin Epi 2001;54:755-65.

129. Rabeneck L, Wristers K, goldstein JL, et al. Reliability, validity, and responsiveness of severity of dyspepsia assessment (SODA) in a randomized clinical trial of a COX-2-specific inhibitor and traditional NSAID therapy. Am J Gastroenterol 2002;97:32-9.

130. Rentz AM, Kahrilas P, Stanghellini V, Tack J, Talley NJ, de la Loge C, Trudeau E, Dubois D, Revicki DA. Development and psychometric evaluation of the patient assessment of upper gastrointestinal symptom severity index (PAGI-SYM) in patients with upper gastrointestinal disorders. Qual Life Res 2004; 13:1737-49.

131. Revicki DA, Rentz AM, Tack J, Stanghellini V, Talley NJ, Kahrilas P, De La Loge C, Trudeau E, Dubois D. Responsiveness and interpretation of a symptom severity index specific to upper gastrointestinal disorders. Clin Gastroenterol Hepatol 2004;2:769-77.

132. Tack J, Van Elzen B, Tytgat G, Wajs E, Van Nueten L, De Ridder F, Boeckxstaens G. A placebo-controlled trial of the 5-HT1A agonist R-137696 on symptoms, visceral hypersensitivity and on impaired accommodation in functional dyspepsia. Neurogastroenterol Motil. 2009 Jun;21(6):619-26, e23-4.

133. De La Loge C, Trudeau E, Marquis P, Revicki DA, Rentz AM, Stanghellini V, Talley NJ, Kahrilas P, Tack J, Dubois D. Responsiveness and interpretation of a quality of life questionnaire specific to upper gastrointestinal disorders. Clin Gastroenterol Hepatol 2004;2:778-86.

134. El-Omar EM, Banerjee S, Wirz A, McColl KEL. The Glasgow dyspepsia severity score-a tool for the global measurement of dyspepsia. Eur J Gastroenterol Hepatol 1996; 8:967-71.

135. McColl K, Murray L, El-Omar E, et al. Symptomatic benefit from eradicating Helicobacter pylori infection in patients with nonulcer dyspepsia. N Eng J Med 1998;339:1869-74

136. Calvet X, Bustamante E, Montserrat A, Roque M, Campo R, Gene B, Brullet E. Validation of phone interview for follow-up in clinical trials on dyspepsia: evaluation of the Glasgow Dyspepsia Severity Score and a Likert-scale symptoms test. Eur J Gasatroenterol Hepatol 2000;12:949-53.

137. Garratt AM, Ruta DA, Russel I, et al. Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. J Clin Epidemiol 1996;49:565-71.

138. Adam B, Liebregts T, Saadat-Gilani K, Vinson B, Holtmann G. Validation of the gastrointestinal symptom score for the assessment of symptoms in patients with functional dyspepsia. Aliment Pharmacol Ther. 2005 Aug 15;22(4):357-63.

139. Madisch A, Holtmann G, Mayr G, Vinson B, Hotz J. Treatment of functional dyspepsia with a herbal preparation. A doubleblind, randomized, placebo-controlled, multicenter trial. Digestion 2004; 69: 45–52.

140. Rosch W, Vinson B, Sassin I. A randomised clinical trial comparing the efficacy of a herbal preparation STW 5 with the prokinetic drug cisapride in patients with dysmotility type of functional dyspepsia. Z Gastroenterol 2002; 40: 401–8.

141. Madisch A, Melderis H, Mayr G, Sassin I, Hotz J. A plant extract and its modified preparation in functional dyspepsia. Results of a double-blind placebo controlled comparative study. Z Gastroenterol 2001; 39: 511–7.

142. Stewart AL, Hays RD, Ware JE Jr. The MOS short-form general health survey. Reliability and validity in a patient population. Med Care 1988;26:724-35.

143. Patrick Dl, Drossman DA, Frederick HO, et al. Quality of life in persons with irritable bowel syndrome. Development and validation of a new measure. Dig Dis Sci 1998;43:400-411.

144. Laine L, Tack J, Howden C, Omar M, Ligozio G, Kralstein J, Earnest DL. Work productivity and daily activity impairment in patients with functional dyspepsia: Double-blind comparison of tegaserod vs placebo. Gastroenterology 2007;132(Suppl. 2):A94.

145. Locke GR, III. Prevalence, incidence and natural history of dyspepsia and functional dyspepsia. In: Talley NJ, ed. Balliere's Clinical Gastroenterology, volume 12, No. 3. London:Balliere Tindall, 1998:435-42.

146. Lindell GH, Celebioglu F, Graffner HO. Non ulcer dyspepsia in the long term perspective. Eur J Gastroenterol Hepatol 1995;7:829-33.

147. Finney JS, Kinnersley N, Hughes M, et al. Meta analysis of antisecretory and gastrokinetic compounds in functional dyspepsia. J Clin Gastroenterol 1998; 26:312-20.

148. Klein KB. Controlled treatment trials in the irritable bowel syndrome: A critique. Gastroenterology 1988;95:232-41.

149. Moayyedi P, Soo S, Deeks J, Delaney B, Innes M, Forman D. Pharmacological interventions for non-ulcer dyspepsia. Cochrane Database of Systematic Reviews 2006;Issue 4:1-22

150. Pappa KA, Williams BO, Payne JE, Buaron KS, Mussari KL, Ciociola AA. A double-blind, placebo-controlled study of the efficacy and safety of non-prescription ranitidine 75 mg in the prevention of meal-induced heartburn. Aliment Pharmacol Ther 1999; 13(4):467-73.

151. FDA Gastrointestinal Drugs Advisory Committee. Transcript from FDA Gastrointestinal Drugs Advisory Committee. Lotronex 1999;November:119.

152. Mangel AW, Hahn BA, Heath AT, et al. Adequate relief as an endpoint in clinical trials in irritable bowel syndrome. J Int Med Res 1998;26:76-81.

153. Mangel AW, Northcutt AR. Review article: safety and efficacy of alosetron, a 5-HT3 receptor antagonist, in female irritable bowel syndrome patients. Aliment Pharmacol Ther 1999;13(Suppl. 2):77-82.

154. Whitehead WE, Palasson OS, Levy RL, Feld AD, VonKorff M, Turner M. Reports of “satisfactory relief” by IBS patients receiving usual medical care are confounded by baseline symptom severity and do not accurately reflect symptom improvement. Am J Gastroenterol 2006;101:1057-65.

155. Corazziari E, Bytzer P, Delvaux M, Holtmann G, Malagelada JR, Morris J, Muller-Lissner S, Spiller RC, Tack J, Whorwell PJ. (2003). Clinical trial guidelines for pharmacological treatment of irritable bowel syndrome. Aliment Pharmacol Ther. 18: 569-80.

TABLES

Table 1. Summary of included trials and outcome measures employed.

Outcome measure	Scoring system	Limitations	References
Binary outcome measure	Adequate/satisfactory relief	Not Rome III-based Does not reflect magnitude of improvement Does not show deterioration May depend on baseline severity	25,40,42,43,53
Individual symptom assessment	(i) Likert scores	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	40,43,44-98,
	(ii) VAS score	Not Rome III-based	99-103
Global evaluation	Likert scale / OTE	Not Rome III-based Recall bias for OTE	42,43, 48, 89, 91, 94-98 70-97,104-107,110,
Disease specific questionnaires	GOS	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	94-96
	GSRS	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	88,89,90,94-97,112,113,122
	Leeds	Not Rome III-based Inclusion of non-FD symptoms Investigator-administered	91,92

		Complex scoring system	
	DSSI	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	42, 117
	Hong Kong Index of Dyspepsia	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	119
	PADYQ	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	120
	Nepean Dyspepsia Index (NDI)	Not Rome III-based Inclusion of non-FD symptoms	42,43,91,92,123,131-135
	SODA	Not Rome III-based Inclusion of non-FD symptoms Does not assess frequency of symptoms	97, 137
	PAGI	Not Rome III-based Inclusion of non-FD symptoms Equal weight to individual symptoms	132
	GDSS	Not Rome III-based Inclusion of non-FD symptoms Investigator-administered Heterogeneous scoring system	135
Quality of life		Not Rome III-based	102,108 (primary

assessments		Not validated as primary outcome measure	outcome measure), 66,70,72,73,78,98-100,103-107 (secondary outcome measures)

I would like to see some further mention of the high placebo response rates in FD (the authors mention this briefly in the Discussion) and the challenge in ascertaining a significant difference between placebo and active management in disease specific measures and overall improvement. How much of a difference is worthwhile? Many researchers are of the view that in functional disorders improvements in specific symptoms are not as important as an overall improvement in well being. Specific symptoms measures may be the wrong trading currency in functional problems, which, by their nature are likely to be part of a global disorder (e.g. tiredness included).

Response: We added a paragraph on the placebo response; the available data show an influence of threshold level choice, but not of the type of endpoint chosen (binary vs overall vs composite). We also, briefly, addressed the (lack of data and consensus on) minimally relevant margin over placebo. We also added a statement on the limitations of addressing only specific FD symptoms in assessment of therapeutic benefit.

.