Systematic review

Received: 10 July 2017 | Accepted: 5 February 2018

DOI: 10.1111/nmo.13327

REVIEW ARTICLE

WILEY Neurogastroenterology & Motility NGM

Systematic review: Disease-specific instruments to assess gastrointestinal symptoms in functional dyspepsia

F. G. M. Smeets[1,2] | A. A. M. Masclee[1,2] | J. M. Conchillo[1,2] | D. Keszthelyi[1,2]

[1]Division of Gastroenterology-Hepatology, Department of Internal Medicine, Maastricht University Medical Center (MUMC+), Maastricht, The Netherlands

[2]NUTRIM, School of Nutrition and Translational Research in Metabolism, Maastricht University, Maastricht, The Netherlands

Correspondence
F. G. M. Smeets, Division of Gastroenterology-Hepatology, Department of Internal Medicine, Maastricht University Medical Center (MUMC+), Maastricht, The Netherlands.
Email: fabienne.smeets@maastrichtuniversity.nl

Abstract

Background: Functional dyspepsia (FD) is a common functional gastrointestinal disorder with incompletely understood pathophysiology and heterogeneous symptom presentation. Assessment of treatment efficacy in FD is a methodological challenge as response to treatment must be assessed primarily by measuring subjective symptoms. Therefore, the use of patient-reported outcome measures (PROMs) is recommended by regulatory authorities to assess gastrointestinal symptoms in clinical trials for FD. In the last decades, a multitude of outcome measures has been developed. However, currently no PROM has been approved by the regulatory authorities, and no consensus has been reached with regard to the most relevant outcome measure in FD.

Purpose: This systematic review discusses the available disease-specific outcome measures for assessment of FD symptoms with psychometric validation properties, strengths, and limitations. Moreover, recommendations for use of current available outcome measures are provided, and potential areas of future research are discussed.

KEYWORDS

functional dyspepsia, functional gastrointestinal disorder, patient-reported outcome measures, questionnaires, symptom assessment

1 | INTRODUCTION

Functional dyspepsia (FD) is a common functional gastrointestinal disorder defined by the presence of symptoms thought to originate from the gastroduodenal region in the absence of any organic, systemic, or metabolic disease that is likely to explain the symptoms.[1] Although the exact pathophysiology of FD is still incompletely understood, it is thought to be multifactorial in origin with involvement of gastroduodenal motor and sensory dysfunction, low-grade inflammation of the duodenal mucosa, and impaired mucosal integrity.[1-4] Dyspeptic patients can present with a variety of symptoms and 4 FD core symptoms have been defined according to the ROME IV criteria, including postprandial fullness, early satiation, epigastric burning, and epigastric pain.[5] However, excessive belching, postprandial nausea, and upper abdominal bloating are considered important additional symptoms.[1,5] Due to the heterogeneity of symptoms in patients with FD, it was hypothesized that different pathophysiological mechanisms could underlie different dyspeptic symptoms. Therefore, subdivision of patients into 2 diagnostic categories was proposed by the ROME IV criteria: (i) postprandial distress syndrome (PDS; presence of postprandial fullness and/or early satiety) and (ii) epigastric pain syndrome (EPS; presence of epigastric pain and/or epigastric burning).[1,6] The rationale for dividing patients into PDS and EPS subgroups was that classification might help to guide therapy, that is, prokinetics for patients with PDS and acid suppressive medication for patients with EPS.[6,7] However, treatment of FD remains unsatisfactory in many patients for several reasons. First, development of novel therapies is hindered by the incompletely understood multifactorial pathophysiology. Moreover, assessment of treatment efficacy in FD is a methodological challenge as response to

treatment must be assessed primarily by measuring subjective symptoms. The Food and Drug Administration (FDA) Guidance recommends the use of well-defined patient-reported outcome measures (PROMs) to assess treatment outcome.[8] Patient-reported outcome measures should be reliable, valid, and responsive with an acceptable recall period.[8] Although a multitude of questionnaires has been developed for assessment of FD symptoms, currently no instrument has undergone all development steps as recommended by the FDA. Therefore, no consensus has been reached with regard to the most relevant outcome measure in FD patients. As a consequence, recent studies used different questionnaires to evaluate treatment efficacy in FD patients, which makes comparison of results difficult and prevents formulation of firm treatment recommendations.[9,10]

The aim of this systematic analysis is to review the published literature on methodology for the assessment of gastrointestinal symptoms in FD patients. We sought to answer the following questions:

1. Which disease-specific questionnaires for assessment of gastrointestinal symptoms in patients with FD are currently available?
2. What are the psychometric properties of these assessment tools, are they eligible for monitoring efficacy of treatment in patients with FD, and which questionnaire can be recommended?

Key Points

- Functional dyspepsia is a symptom-based diagnosis according to ROME IV criteria, and use of patient-reported outcome measures is recommended to assess gastrointestinal symptoms in clinical trials.
- Twenty outcome measures were identified to be suitable for assessment of dyspeptic symptoms with varying psychometric validation characteristics. The *Leuven Postprandial Distress Scale* and *Functional Dyspepsia Symptom Diary* seem most promising for use in current clinical trials.
- In the future, limitations of retrospective questionnaires, including recall bias and ecological bias, and noncompliance might potentially be reduced by the use of techniques such as the Experience Sampling Method.

2 | METHODS

2.1 | Search strategy

To identify relevant articles, a structured search in the PubMed, Cochrane, and Embase databases was performed with the following search strategy: (dyspepsia) AND (questionnaire OR outcome measure OR symptom assessment). English was used as filter and articles were selected between June 1970 and September 2017. The reference list of selected articles was examined to identify additional relevant citations.

Title and abstract of all articles were screened by 2 independent researchers (F.S. and D.K.) to assess eligibility according to predefined criteria. Full-text review was performed for studies reporting on psychometric validation properties (eg, reproducibility, validity, and responsiveness) of disease-specific PROMs for patients with (functional) dyspepsia. Although we intended to include only disease-specific outcome measures, frequently used but non-specific questionnaires were included to provide a complete overview and thereby give accurate recommendations for clinical practice. Studies were excluded if they were reviews, comments, or replies on original articles, or abstracts without available full text (presented at conferences). Moreover, studies with another aim (eg, questionnaires used to diagnose dyspepsia) or studies evaluating only quality of life were excluded. Both reviewers agreed on inclusion of the articles. In case of disagreement, a third reviewer decided with regard to the inclusion (A.M.).

2.2 | Outcome measures

The outcomes of interest with regard to this systematic review were (i) the available disease-specific PROMs used to assess gastrointestinal symptoms in FD patients, and (ii) the psychometric properties, strengths and limitations of these PROMs.

In clinical trials with FD patients, several types of disease-specific outcome measures have been used. Disease-specific instruments can be divided into unidimensional (ie, evaluating gastrointestinal symptoms) or multidimensional (ie, evaluating gastrointestinal symptoms and other domains such as quality of life or impact of symptoms on emotional functioning) instruments.[11]

Psychometric validation of a PROM incorporates assessment of reliability, validity, and responsiveness to change.[8] Reliability refers to the ability of a PRO instrument to yield consistent, reproducible results and is assessed by the test-retest reliability (ie, indicates the stability of scores over time when no change in health status is expected) and internal consistency (ie, indicates the correlation between different items comprising a scale of the instrument). Test-retest reliability is often expressed by the intraclass correlation coefficient (ICC) and an ICC above 0.75 is considered adequate.[12-14] Cronbach's α is frequently used to describe internal consistency, and a value between 0.7 and 0.9 is considered acceptable.[8,13,14] Validity can be divided into different subtypes. Content validity describes the extent to which the instrument measures all symptoms indicative of FD, whereas construct validity provides evidence for the degree of correlation between related instruments measuring the same symptoms. Construct validity is expressed by the Pearson or Spearman correlation coefficient referred to as *r*, and a correlation coefficient between 0.4 and 0.8 is considered appropriate.[8] Known groups validity reflects the ability of a PROM to differentiate between groups which would theoretically be expected to differ in

symptom scores. Responsiveness of the instrument is the ability to detect clinically meaningful changes in symptoms over time. Several methods are used for assessment of responsiveness. First, the (significant) change in symptom scores after treatment can be assessed. Moreover, effect sizes (ES) and standardized response means (SRMs) can be measured. The effect size (ES) estimates the change in symptom score (δ) relative to the variability of the symptom score at baseline (ie, ES = δ symptom score/standard deviation symptom score at baseline). Outcome measures with ES > 0.8 are considered responsive to change.[15] The SRM is calculated by dividing the change in symptom score (δ) between the standard deviation of the change (ie, SRM = δ symptom score/standard deviation of the change). Outcome measures with SRM > 0.8 are considered responsive to change.[16]

Besides psychometric validation properties, we evaluated the following characteristics of PROMs for assessment of strengths and limitations: (i) whether validation was performed in the target population (ie, patients with FD according to ROME criteria), (ii) recall period, (iii) assessment of FD core symptoms (ROME IV; ie, postprandial fullness, early satiation, epigastric pain, and epigastric burning) and additional dyspeptic symptoms (ie, nausea, bloating, and belching), and (iv) available minimum clinically important difference (MCID).

Data extraction was performed by 2 independent researchers (F.S. and D.K.). In case of disagreement, differences were discussed and resolved.

3 | RESULTS

The systemic literature search and screening of reference lists of selected articles identified 4304 citations after removal of duplicates. After abstract and full-text review, 40 articles were included in this systematic review for analysis of disease-specific outcome measures (Figure 1). The disease-specific outcome measures included in this review are divided into 2 categories: (i) unidimensional measures (n = 14) and (ii) multidimensional measures (n = 6). Table 1 provides a short overview of the available outcome measures. Extensive descriptions of all individual questionnaires are available as Supplementary Data. In short, we summarize here the following aspects examined.

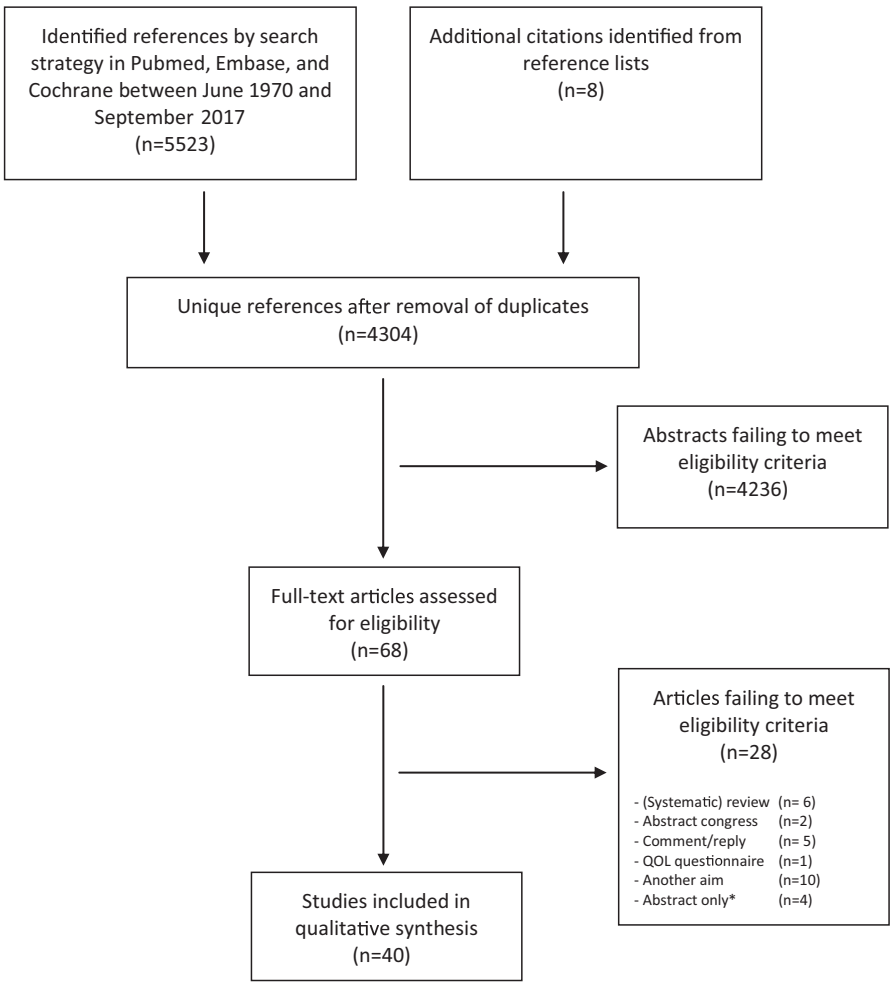

FIGURE 1 Flow diagram showing the studies included in the systematic literature review

QOL = quality of life questionnaire
*Abstract only: full text not available after contacting corresponding author and/or journal

TABLE 1 Overview of available patient-reported outcome measures

Assessment method	Unidimensional or multidimensional PROM	Number of items	Recall period	Investigator vs patient administered
Dyspepsia Symptom Severity Index (DSSI)	Unidimensional	20	2 weeks	Patient
Global Overall Symptom Scale (GOS)	Unidimensional	10	2 days (GOS2) or 28 days (GOS28)	Patient
Leeds Dyspepsia Questionnaire (LDQ)	Unidimensional	8 (and 1 item to assess the most troublesome symptom)	6 months	Investigator
Short-form Leeds Dyspepsia Questionnaire	Unidimensional	4 (and 1 item to assess the most troublesome symptom)	2 months	Patient
Gastrointestinal Symptom Rating Scale (GSRS)	Unidimensional	15	1 week or 1 month	Patient
Hong Kong Dyspepsia Index	Unidimensional	12	Undefined	Patient
Severity Index of Bologna	Unidimensional	8	Undefined	Investigator
Porto Alegre Dyspeptic Symptoms Questionnaire (PADYQ)	Unidimensional	11	1 month	Investigator
Gastrointestinal Symptom Score	Unidimensional	10	Undefined	Investigator
Gastrointestinal Symptom Severity Index	Unidimensional	39	14 days or 30 days	Patient
Digestive Health Status Instrument (DHSI)	Unidimensional	34	1 month	Patient
Leuven Postprandial Distress Scale (LPDS)	Unidimensional	8	1 day	Patient
Functional Dyspepsia Symptom Diary (FDSD)	Unidimensional	8	1 day	Patient
Patient Assessment of Upper Gastrointestinal Symptom Severity Index (PAGI-SYM)	Unidimensional	20	2 weeks	Patient
Nepean Dyspepsia Index (NDI)	Multidimensional	Quality of life component: 25 Symptom checklist: 15	2 weeks	Patient
Short-form Nepean Dyspepsia Index (SF-NDI)	Multidimensional	10	2 weeks	Patient
Glasgow Dyspepsia Severity Index (GDSS)	Multidimensional	8	3 or 6 months	Investigator
Spanish Language Dyspepsia Questionnaire (SLDQ)	Multidimensional	38	Undefined	Investigator
The Clinical Dyspepsia Questionnaire (CDQ)	Multidimensional	15	Undefined	Patient
Severity of Dyspepsia Assessment (SODA)	Multidimensional	17	1 day or 1 week	Patient

3.1 | Reliability

Internal consistency and/or test-retest reliability were assessed in the majority of questionnaires with moderate to acceptable results (Cronbach's α ranging from 0.43 to 0.96; intraclass correlation coefficient (ICC) ranging from 0.42 to 0.96). Test-retest reliability was considered too low for the Global Overall Symptom Scale, several domains of the Gastrointestinal Symptom Rating Scale, and Severity Index of Bologna. Four outcome measures assessed only internal consistency without test-retest reliability, including the Digestive Health Status Instrument, Short-Form Nepean Dyspepsia Index, Functional Dyspepsia Symptom Diary, and Severity of Dyspepsia Assessment.[17-21] As test-retest reliability is considered the most important type of reliability for PRO instruments used in clinical trials, there is insufficient evidence for adequate reliability of these 4 outcome measures.[8] For several investigator-administered questionnaires including the Leeds Dyspepsia Questionnaire and Glasgow Dyspepsia Severity Index, both intra-rater and interrater reliability were assessed and showed good to excellent results.[22,23]

3.2 | Validity

Content validity was assessed by 10 of the described questionnaires and was considered good by either experts in the field or by patients during focus groups or cognitive interviews. Moreover, good known-groups validity was described in 12 of the questionnaires including the recently developed Functional Dyspepsia

Symptom Diary. Widely varying results with regard to construct validity have been described for individual questionnaires, although the majority was correlated with subscales of either the Short-Form 36 Health Survey or the Patient Assessment of Upper Gastrointestinal Symptom Severity Index.

One commonly used questionnaire is the Nepean Dyspepsia Index which was developed as disease-specific quality of life (QOL) measure but also contains a symptom checklist for assessment of gastrointestinal symptoms.[24] Although known-groups validity has been assessed for both components of the Nepean Dyspepsia Index, construct validity has not been assessed for the symptom checklist. Moreover, construct validity of the Short-Form Nepean Dyspepsia Index was only assessed by correlating scores with the Nepean Dyspepsia Index but not with independent gastrointestinal questionnaires.[20]

3.3 | Responsiveness

Assessment of responsiveness is considered essential for use of outcome measures in clinical trials. Although responsiveness was adequate for the majority of questionnaires, it was not (yet) assessed for 5 outcome measures (ie, Functional Dyspepsia Symptom Diary, Dyspepsia Symptom Severity Index, Gastrointestinal Symptom Severity Index, Spanish Language Dyspepsia Questionnaire, and The Clinical Dyspepsia Questionnaire), which limits the use of these outcome measures in clinical trials.[17-19,25-28]

The Leuven Postprandial Distress Scale is a recently developed outcome measure for FD patients primarily with the PDS subtype. Assessment of responsiveness was performed in 60 patients according to the ROME III criteria (PDS: 48% and PDS-EPS overlap: 52%). Whereas responsiveness was demonstrated for patients with PDS symptoms, no correlation was found between the overall treatment evaluation and change of symptom scores for patients with EPS symptoms. This might be due to the more prevalent and more severe PDS symptoms of included patients.

3.4 | Additional characteristics

3.4.1 | Validation in target population

Of the available PROMs, only the Leuven Postprandial Distress Scale and Functional Dyspepsia Symptom Diary have been validated in FD patients according to ROME III criteria.

3.4.2 | Investigator vs patient administration

Six outcome measures are investigator-administered instead of patient-administered (ie, Leeds Dyspepsia Questionnaire, Severity Index of Bologna, Porto Alegre Dyspeptic Symptoms Questionnaire, Gastrointestinal Symptom Score, Glasgow Dyspepsia Severity Index, and Spanish Language Dyspepsia Questionnaire).

3.4.3 | Recall period

The recall period ranges between 1 day and 6 months, whereas the recall period has been undefined for 5 PROMs (ie, Hong Kong Dyspepsia Index, Severity Index of Bologna, Gastrointestinal Symptom Score, Spanish Language Dyspepsia Questionnaire, and The Clinical Dyspepsia Questionnaire). Although no optimal recall period has been defined for FD,[5,8,29] the FDA recommends daily assessment of symptoms in functional gastrointestinal disorders due to symptom variability.[5,29] Only the Leuven Postprandial Distress Scale, Functional Dyspepsia Symptom Diary, and modified Severity of Dyspepsia Assessment (mSODA) have a recall period of 24 hours.

3.4.4 | Core symptoms of functional dyspepsia

Assessment of all 4 FD core symptoms according to the ROME IV criteria is limited to 4 PROMs (ie, Dyspepsia Symptom Severity Index, Leuven Postprandial Distress Scale, Symptom Checklist of the Nepean Dyspepsia Index, and Functional Dyspepsia Symptom Diary).

3.4.5 | Minimum clinically important difference (MCID)

A minimum clinically important difference (MCID) is available for 5 outcome measures (ie, Gastrointestinal Symptom Rating Scale, Leuven Postprandial Distress Scale, Patient Assessment of Upper Gastrointestinal Symptom Severity Index, Nepean Dyspepsia Index, and modified Severity of Dyspepsia Assessment).

4 | DISCUSSION

In this systematic review, the currently available disease-specific instruments for assessment of symptoms in patients with FD have been evaluated. Although the development of a PROM is necessary to uniformly assess FD symptoms in clinical trials for evaluation of treatment efficacy, no universally accepted PROM is currently available.

4.1 | Development of PROMs

Development of a PROM is a comprehensive process and consists of several steps. Regulatory authorities (FDA) suggest performing a literature review and obtaining expert opinion, followed by patient focus groups and cognitive debriefing. Input from the target population is essential for item generation, and patient focus group interviews were performed in the development of several PROMs (eg, Leuven Postprandial Distress Scale, Functional Dyspepsia Symptom Diary, Dyspepsia Symptom Severity Index, and Nepean Dyspepsia Index).

4.2 | Psychometric validation of PROMS

Psychometric validation has to be performed evaluating reliability, validity, and responsiveness.[8] The majority of the described

questionnaires proved to be reliable and valid. The lower test-retest reliability for the abdominal pain domain of the Gastrointestinal Symptom Rating Scale might be due to the fact that this domain includes only 2 items. One also has to take into account the complex assessment of pain due to symptom variability. For instance, for the Global Overall Symptom Scale-28, test-retest reliability was assessed with an interval of 56 days which might be too long for accurate assessment of "stable" gastrointestinal symptoms.

In order to evaluate treatment efficacy in clinical trials, a PROM should be responsive to change. Responsiveness of the Leuven Postprandial Distress Scale has only been demonstrated for PDS symptoms, not for EPS.[30] Moreover, psychometric validation has to be performed in a population of FD patients, and currently, only the Leuven Postprandial Distress Scale and Functional Dyspepsia Symptom Diary have been evaluated in FD patients according to the Rome III criteria.[8,19,29,30] Therefore, the majority of the instruments we evaluated do not fulfill all psychometric validation criteria as defined by the FDA.[31]

4.3 | Patient vs investigator-administered PROMs

The FDA defines a PROM as any report of the status of a patient's health condition that comes directly from the patient without interpretation of the patient's response by a clinician.[8] Patient self-completion of a PROM is recommended to reduce social desirability bias.[11] In addition, recent studies described poor to moderate overall agreement between patients' and expert clinicians' rating of symptom severity in FD patients, and underestimation of several symptoms (eg, early satiety and abdominal pain) by clinicians which favors the use of a patient-administered outcome measure.[19,32,33] This might limit the use of the Leeds Dyspepsia Questionnaire, Porto Alegre Dyspeptic Symptoms Questionnaire, Gastrointestinal Symptom Score, and Glasgow Dyspepsia Severity Index as these are investigator-administered instruments.[22,23,34,35]

4.4 | Functional dyspepsia: core and additional symptoms

Another challenge for development of a PROM is the multifactorial nature of FD and the diverse symptom presentation of patients. As the FDA states that in functional disorders the effect of treatment should be measured on each symptom in order to ensure that treatment does not negatively affect symptoms, we suggest that a PROM should evaluate at least the 4 FD core symptoms (early satiation, postprandial fullness, epigastric burning, and epigastric pain[5]) and the 3 additional symptoms (upper abdominal bloating, postprandial nausea, and excessive belching).[8,36] The latter 3 were reported by a substantial subgroup of FD patients during patient focus group interviews, cognitive interviews, and validation studies (ie, bloating 33%-93%, nausea 31%-73%, belching 27%-69%).[1,5,30,37] Interestingly, during development of the Functional Dyspepsia Symptom Diary, dyspeptic patients ranked "bloating"—which is not a core symptom according to Rome IV—as the most important symptom for improvement with effective therapy.[37] Consequently, the total symptom score of the Functional Dyspepsia Symptom Diary consists of the scores for the 4 Rome IV core FD symptoms and for bloating, with the other 2 additional symptoms of belching and nausea being evaluated separately as individual items. Currently, no quantitative data are available for direct comparison of the validity, reliability, and responsiveness with assessment of 4 vs 5 or 7 FD symptoms by PROMs. Therefore, until such data become available, we favor assessment of the 4 core symptoms and the 3 additional symptoms.

4.5 | Limitations of retrospective outcome measures

For accurate assessment of FD core symptoms, an accurate recall period should be used to prevent recall bias as the available retrospective questionnaires are based on autobiographical memory.[38] Only the Leuven Postprandial Distress Scale, Functional Dyspepsia Symptom Diary, and modified Severity of Dyspepsia Assessment have a recall period of 24 hours, in accordance with the FDA recommendation.

Besides recall bias, 2 additional limitations of retrospective questionnaires should be considered. First, symptom variability in functional disorders might be due to the influence of triggers and circumstances. For example, FD symptoms might be influenced by intake of food (eg, fat) and psychological factors (eg, anxiety).[39-41] Therefore, ecological bias can occur when questionnaires are completed in another environment or situation compared with situations in which symptoms were triggered. Furthermore, the lack of patient compliance is another major limitation of paper questionnaires. This limitation can potentially be eliminated by the use of electronic instruments.[42]

4.6 | Recommendation for the use of current available instruments

In 2011, Ang et al. published an extensive review about PROMs for FD. None of the available PROMs was considered appropriate for use in clinical trials which warranted the urgent need for development and validation of novel outcome measures. Since 2011, new PROMs have become available (ie, Leuven Postprandial Distress Scale, Functional Dyspepsia Symptom Diary, Gastrointestinal Symptom Severity Index), which have been included in the present systematic review. However, neither an universally validated and accepted patient-reported outcome measure for assessment of treatment efficacy in clinical trials of FD patients, nor a FD-specific recommendation of the FDA are currently available. Therefore, recent clinical trials used different questionnaires and methods to assess treatment efficacy, including the Nepean Dyspepsia Index, daily symptom diaries, a binary outcome measure (relief or no relief of symptoms), PAGI-SYM, and assessment of overall treatment effect.[9,10,43-45] Therefore, selection of one or a limited number of current available outcome measures is necessary to uniformly assess FD symptoms in clinical trials. The question arises which questionnaires should be

recommended based on the results of the current systematic review, pending the proper validation of novel PROMs.

The Leuven Postprandial Distress Scale and Functional Dyspepsia Symptom Diary appear to be promising outcome measures in FD patients as item selection was performed by patient focus groups and/or cognitive debriefing, the questionnaires are patient-completed, contain all FD core and additional dyspeptic symptoms, and have a short recall period of 24 hours. For the recently developed Functional Dyspepsia Symptom Diary, only preliminary data are available with regard to psychometric validation. Although good internal consistency and known-groups validity were described, additional research is necessary to evaluate test-retest reliability, construct validity, and responsiveness before use of this PROM in clinical trials. The Leuven Postprandial Distress Scale demonstrated good reliability, validity, and responsiveness for the PDS domain. Additional studies are necessary to validate use of the Leuven Postprandial Distress Scale in patients with EPS predominant symptoms. In the case that psychometric validation for EPS symptoms is confirmed for (a modified version of) the Leuven Postprandial Distress Scale and additional psychometric validation is performed for the Functional Dyspepsia Symptom Diary, these PROMs might prove to be excellent outcome measures to evaluate the full range of FD symptoms, as recommended by the FDA.

The use of a single universal questionnaire for assessment of the full range of FD symptoms seems favorable for several reasons. First, previous studies demonstrated a high percentage of patients fulfilling both criteria for EPS and PDS ranging from 34.4% to 66%.[30,46-49] In addition, the subdivision of patients into PDS and EPS subtypes was proposed to serve as guide for diagnostic and therapeutic approaches as it was hypothesized that distinct pathophysiological processes were present between the 2 subtypes. Current evidence for the presence of distinct etiopathogenesis is limited and inconclusive. Previous studies found no difference in gastric accommodation between the 2 FD subtypes, and contradictory results were obtained with regard to gastric hypersensitivity and gastric emptying.[50,51] Division of FD patients into subtypes does currently not reliably distinguish subgroups with a different underlying pathophysiology. The use of this classification to guide therapy is, therefore, limited and favors the use of a single universal PROM for assessment of all FD core symptoms.[2,52-55]

Apart from the Leuven Postprandial Distress Scale and Functional Dyspepsia Symptom Diary, 2 questionnaires assess all FD core symptoms: (i) the symptom checklist of the Nepean Dyspepsia Index and (ii) the Dyspepsia Symptom Severity Index.[5] Use of the Dyspepsia Symptom Severity Index in clinical trials for the evaluation of treatment efficacy is limited by the fact that responsiveness has never been evaluated.[25] The Nepean Dyspepsia Index has widely been used and is considered a valid, reliable, responsive questionnaire for assessment of FD symptoms with a recall period of 2 weeks. However, it is important to consider that the Nepean Dyspepsia Index measures quality of life, whereas the symptom checklist assesses the severity and frequency of dyspeptic symptoms. Although psychometric validation has been performed for the Nepean Dyspepsia Index, test-retest reliability and convergent validity have never been assessed for the separate symptom checklist.[20,24,56,57]

5 | FUTURE PERSPECTIVES

5.1 | Patient-reported outcome measurement system

In 2004, the National Institute of Health (NIH) has launched the patient-reported outcome measurement system (PROMIS) to develop a set of short, computer-based, reliable PROMs to assess physical, mental, and social health for research purposes and clinical care. Short outcome measures for 8 gastrointestinal symptom complexes were developed with good reliability and validity.[58,59] Although the use of these domain-specific PROMs was hypothesized to improve patient satisfaction, a recently performed trial compared the use of PROMIS with regular care and found no difference in patient satisfaction and shared decision making.[60] However, several limitations should be considered including the poor response rate, lack of information about influence on clinic visit efficiency and frequency, cost-effectiveness, and long-term effects, which demonstrates the need for additional research.[60,61]

5.2 | Experience sampling method

Real-time assessment of symptoms is recommended to minimize recall bias, and the use of an e-health approach seems favorable to improve patient comfort and reduce non-compliance. The experience sampling method (ESM) is a digital method used for "real-time symptom assessment" which has been validated in psychological studies.[62] Repeated measures are used to assess variability in symptoms and might provide information about potential triggers as it takes into account situations and circumstances (eg, meal intake, social environment, psychological factors) that have occurred prior to symptom generation which might reduce ecological bias. Previous studies described associations between FD, anxiety, and depression. Assessment of psychological factors, besides gastrointestinal symptoms, might, therefore, provide insight into psychological triggers with potential treatment consequences. Inclusion of pictograms, besides verbal descriptors, might be another advantage of the digital ESM. Recently, Tack et al. developed a set of pictograms to accompany verbal descriptors for dyspeptic symptoms.[19] Use of pictograms enhanced concordance of FD symptom rating by patients compared with physicians' symptom evaluation.[19] Addition of pictograms might, therefore, be an useful adjunctive for the development of an ESM-based PROM.[29] One potential relevant limitation of the ESM might be a higher patient burden due to the repeated measures at random moments during daily life. Use of a restricted number of questions might reduce this limitation.[63]

Use of the ESM in patients with gastrointestinal diseases is currently limited. Two studies evaluated the use of ESM in IBS patients and demonstrated that patients reported higher scores for

abdominal pain in retrospective questionnaires compared with mean scores derived with ESM.[64,65] To our knowledge, the use of the ESM in patients with FD has not been described previously. Therefore, extensive item selection and psychometric validation of the ESM in FD patients has to be performed.

6 | CONCLUSION

In conclusion, multiple different questionnaires are currently used in clinical trials to assess gastrointestinal symptoms in patients with FD. Due to the lack of a validated and widely accepted PROM, results of clinical trials are difficult to compare demonstrating the need for a universal PROM. Although current available questionnaires do not fulfill all criteria for psychometric validation, the *Leuven Postprandial Distress Scale* and *Functional Dyspepsia Symptom Diary* seem most promising. However, retrospective questionnaires have several limitations that might potentially be reduced by the use of novel digital approaches such as the ESM, which seems a promising method for use in patients with functional gastrointestinal disorders.

ACKNOWLEDGMENT

None.

CONFLICTS OF INTEREST

None for all authors.

AUTHOR CONTRIBUTION

FGMS study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript; AAMM study concept and design, analysis and interpretation of data, critical revision of the manuscript for important intellectual content; JMC analysis and interpretation of data, critical revision of the manuscript for important intellectual content; DK study concept and design, acquisition of data, analysis and interpretation of data, drafting of the manuscript, critical revision of the manuscript for important intellectual content. All authors have reviewed and approved the final manuscript.

ORCID

F. G. M. Smeets http://orcid.org/0000-0002-1761-8312

REFERENCES

1. Stanghellini V, Chan FK, Hasler WL, et al. Gastroduodenal disorders. *Gastroenterology*. 2016;150:1380-1392.
2. Vanheel H, Vicario M, Vanuytsel T, et al. Impaired duodenal mucosal integrity and low-grade inflammation in functional dyspepsia. *Gut*. 2014;63:262-271.
3. Keely S, Walker MM, Marks E, Talley NJ. Immune dysregulation in the functional gastrointestinal disorders. *Eur J Clin Invest*. 2015;45:1350-1359.
4. Carbone F, Tack J. Gastroduodenal mechanisms underlying functional gastric disorders. *Dig Dis*. 2014;32:222-229.
5. Taylor F, Reasner DS, Carson RT, et al. Development of a symptom-based patient-reported outcome instrument for functional dyspepsia: a preliminary conceptual model and an evaluation of the adequacy of existing instruments. *Patient*. 2016;9:409-418.
6. Talley NJ, Ford AC. Functional dyspepsia. *N Engl J Med*. 2015;373:1853-1863.
7. Vanheel H, Tack J. Therapeutic options for functional dyspepsia. *Dig Dis*. 2014;32:230-234.
8. FaDA (FDA). Guidance for Industry; Patient-Reported Outcome Measures: Use in Medical Product Development to Support Labeling Claims. 2009.
9. Talley NJ, Locke GR, Saito YA, et al. Effect of amitriptyline and escitalopram on functional dyspepsia: a multicenter, randomized controlled study. *Gastroenterology*. 2015;149:340-349 e2.
10. Braak B, Klooker TK, Wouters MM, Lei A, van den Wijngaard RM, Boeckxstaens GE. Randomised clinical trial: the effects of amitriptyline on drinking capacity and symptoms in patients with functional dyspepsia, a double-blind placebo-controlled study. *Aliment Pharmacol Ther*. 2011;34:638-648.
11. Fraser A, Delaney B, Moayyedi P. Symptom-based outcome measures for dyspepsia and GERD trials: a systematic review. *Am J Gastroenterol*. 2005;100:442-452.
12. Streiner DL, Norman GR. *Health Measurement Scales: A Practical Guide to Their Development and Use*. New York, NY: Oxford University Press; 2003.
13. Alrubaiy L, Hutchings HA, Williams JG. Assessing patient reported outcome measures: a practical guide for gastroenterologists. *United European Gastroenterol J*. 2014;2:463-470.
14. Terwee CB, Bot SD, de Boer MR, et al. Quality criteria were proposed for measurement properties of health status questionnaires. *J Clin Epidemiol*. 2007;60:34-42.
15. Talley NJ, Fullerton S, Junghard O, Wiklund I. Quality of life in patients with endoscopy-negative heartburn: reliability and sensitivity of disease-specific instruments. *Am J Gastroenterol*. 2001;96:1998-2004.
16. Berglund M, Thelander S, Salaspuro M, Franck J, Andreasson S, Ojehagen A. Treatment of alcohol abuse: an evidence-based review. *Alcohol Clin Exp Res*. 2003;27:1645-1656.
17. Shaw M, Talley NJ, Adlis S, Beebe T, Tomshine P, Healey M. Development of a digestive health status instrument: tests of scaling assumptions, structure and reliability in a primary care population. *Aliment Pharmacol Ther*. 1998;12:1067-1078.
18. Shaw MJ, Beebe TJ, Adlis SA, Talley NJ. Reliability and validity of the digestive health status instrument in samples of community, primary care, and gastroenterology patients. *Aliment Pharmacol Ther*. 2001;15:981-987.
19. Tack J, Carbone F, Holvoet L, Vanheel H, Vanuytsel T, Vandenberghe A. The use of pictograms improves symptom evaluation by patients with functional dyspepsia. *Aliment Pharmacol Ther*. 2014;40:523-530.
20. 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;15:207-216.
21. Rabeneck L, Cook KF, Wristers K, Souchek J, Menke T, Wray NP. SODA (severity of dyspepsia assessment): a new effective outcome measure for dyspepsia-related health. *J Clin Epidemiol*. 2001;54:755-765.
22. Moayyedi P, Duffett S, Braunholtz D, et al. The Leeds Dyspepsia Questionnaire: a valid tool for measuring the presence and severity of dyspepsia. *Aliment Pharmacol Ther*. 1998;12:1257-1262.

23. el-Omar EM, Banerjee S, Wirz A, McColl KE. The Glasgow Dyspepsia Severity Score–a tool for the global measurement of dyspepsia. *Eur J Gastro Hepatol*. 1996;8:967-971.
24. Talley NJ, Haque M, Wyeth JW, et al. Development of a new dyspepsia impact scale: the Nepean Dyspepsia Index. *Aliment Pharmacol Ther*. 1999;13:225-235.
25. Leidy NK, Farup C, Rentz AM, Ganoczy D, Koch KL. Patient-based assessment in dyspepsia: development and validation of Dyspepsia Symptom Severity Index (DSSI). *Dig Dis Sci*. 2000;45:1172-1179.
26. Crowell MD, Umar SB, Lacy BE, Jones MP, DiBaise JK, Talley NJ. Multi-dimensional gastrointestinal symptom severity index: validation of a brief GI symptom assessment tool. *Dig Dis Sci*. 2015;60:2270-2279.
27. Goldman J, Conrad DF, Ley C, et al. Validation of Spanish language dyspepsia questionnaire. *Dig Dis Sci*. 2002;47:624-640.
28. Garratt AM, Ruta DA, Russell I, et al. Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. *J Clin Epidemiol*. 1996;49:565-571.
29. Irvine EJ, Tack J, Crowell MD, et al. Design of treatment trials for functional gastrointestinal disorders. *Gastroenterology*. 2016;150:1469-1480 e1.
30. Carbone F, Vandenberghe A, Holvoet L, et al. Validation of the Leuven Postprandial Distress Scale, a questionnaire for symptom assessment in the functional dyspepsia/postprandial distress syndrome. *Aliment Pharmacol Ther*. 2016;44:989-1001.
31. Ang D, Talley NJ, Simren M, Janssen P, Boeckxstaens G, Tack J. Review article: endpoints used in functional dyspepsia drug therapy trials. *Aliment Pharmacol Ther*. 2011;33:634-649.
32. Fransen GA, Janssen MJ, Muris JW, Mesters I, Knottnerus JA. Measuring the severity of upper gastrointestinal complaints: does GP assessment correspond with patients' self-assessment? *Fam Pract*. 2007;24:252-258.
33. Madsen LG, Hansen JM, Gronvold M, Bytzer P. The validity of a symptom diary in ratings of dyspepsia measured against a detailed interview: do patients and clinicians agree in their assessment of symptoms? *Aliment Pharmacol Ther*. 2007;26:905-912.
34. 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;49:1822-1829.
35. 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;22:357-363.
36. Administration FaD. Gastroparesis: Clinical Evaluation of Drugs for Treatment; Guidance for Industry. 2015.
37. Taylor F, Higgins S, Carson RT, et al. Development of a symptom-focused patient-reported outcome measure for functional dyspepsia: the functional dyspepsia symptom diary (FDSD). *Am J Gastroenterol*. 2018;113:39-48.
38. Bradburn NM, Rips LJ, Shevell SK. Answering autobiographical questions: the impact of memory and inference on surveys. *Science*. 1987;236:157-161.
39. Feinle-Bisset C, Horowitz M. Dietary factors in functional dyspepsia. *Neurogastroenterol Motil*. 2006;18:608-618.
40. Aro P, Talley NJ, Johansson SE, Agreus L, Ronkainen J. Anxiety is linked to new-onset dyspepsia in the Swedish population: a 10-year follow-up study. *Gastroenterology*. 2015;148:928-937.
41. Mak AD, Wu JC, Chan Y, Chan FK, Sung JJ, Lee S. Dyspepsia is strongly associated with major depression and generalised anxiety disorder – a community study. *Aliment Pharmacol Ther*. 2012;36:800-810.
42. Stone AA, Shiffman S, Schwartz JE, Broderick JE, Hufford MR. Patient compliance with paper and electronic diaries. *Control Clin Trials*. 2003;24:182-199.
43. Tack J, Ly HG, Carbone F, et al. Efficacy of mirtazapine in patients with functional dyspepsia and weight loss. *Clin Gastroenterol Hepatol*. 2016;14:385-392 e4.
44. van Kerkhoven LA, Laheij RJ, Aparicio N, et al. Effect of the antidepressant venlafaxine in functional dyspepsia: a randomized, double-blind, placebo-controlled trial. *Clin Gastroenterol Hepatol*. 2008;6:746-752.
45. Matsueda K, Hongo M, Tack J, Saito Y, Kato H. A placebo-controlled trial of acotiamide for meal-related symptoms of functional dyspepsia. *Gut*. 2012;61:821-828.
46. Carbone F, Holvoet L, Tack J. Rome III functional dyspepsia subdivision in PDS and EPS: recognizing postprandial symptoms reduces overlap. *Neurogastroenterol Motil*. 2015;27:1069-1074.
47. Fang YJ, Liou JM, Chen CC, et al. Distinct aetiopathogenesis in subgroups of functional dyspepsia according to the Rome III criteria. *Gut*. 2015;64:1517-1528.
48. Vakil N, Halling K, Ohlsson L, Wernersson B. Symptom overlap between postprandial distress and epigastric pain syndromes of the Rome III dyspepsia classification. *Am J Gastroenterol*. 2013;108:767-774.
49. Carbone F, Holvoet L, Vanuytsel T, Tack J. Rome III functional dyspepsia symptoms classification: severity vs frequency. *Neurogastroenterol Motil*. 2017;29. https://doi.org/10.1111/nmo.13024.
50. Vanheel H, Carbone F, Valvekens L, et al. Pathophysiological abnormalities in functional dyspepsia subgroups according to the Rome III criteria. *Am J Gastroenterol*. 2017;112:132-140.
51. Di Stefano M, Miceli E, Tana P, et al. Fasting and postprandial gastric sensorimotor activity in functional dyspepsia: postprandial distress vs. epigastric pain syndrome. *Am J Gastroenterol*. 2014;109:1631-1639.
52. Talley NJ. Editorial: moving away from focussing on gastric pathophysiology in functional dyspepsia: new insights and therapeutic implications. *Am J Gastroenterol*. 2017;112:141-144.
53. Lee KJ, Tack J. Duodenal implications in the pathophysiology of functional dyspepsia. *J Neurogastroenterol Motil*. 2010;16:251-257.
54. Walker MM, Aggarwal KR, Shim LS, et al. Duodenal eosinophilia and early satiety in functional dyspepsia: confirmation of a positive association in an Australian cohort. *J Gastroenterol Hepatol*. 2014;29:474-479.
55. Liebregts T, Adam B, Bredack C, et al. Small bowel homing T cells are associated with symptoms and delayed gastric emptying in functional dyspepsia. *Am J Gastroenterol*. 2011;106:1089-1098.
56. Talley NJ, Verlinden M, Jones M. 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;94:2390-2397.
57. Jones M, Talley NJ. Minimum clinically important difference for the Nepean Dyspepsia Index, a validated quality of life scale for functional dyspepsia. *Am J Gastroenterol*. 2009;104:1483-1488.
58. Khanna P, Agarwal N, Khanna D, et al. Development of an online library of patient-reported outcome measures in gastroenterology: the GI-PRO database. *Am J Gastroenterol*. 2014;109:234-248.
59. Spiegel BM, Hays RD, Bolus R, et al. Development of the NIH patient-reported outcomes measurement information system (PROMIS) gastrointestinal symptom scales. *Am J Gastroenterol*. 2014;109:1804-1814.
60. Almario CV, Chey WD, Khanna D, et al. Impact of National Institutes of Health Gastrointestinal PROMIS measures in clinical practice: results of a multicenter controlled trial. *Am J Gastroenterol*. 2016;111:1546-1556.
61. Gracie DJ, Ford AC. Editorial: using patient-reported outcome measures in gastroenterology: PROMISed land or road to nowhere? *Am J Gastroenterol*. 2016;111:1557-1558.

62. Myin-Germeys I, Oorschot M, Collip D, Lataster J, Delespaul P, van Os J. Experience sampling research in psychopathology: opening the black box of daily life. *Psychol Med*. 2009;39:1533-1547.
63. Stone AA, Broderick JE, Schwartz JE, Shiffman S, Litcher-Kelly L, Calvanese P. Intensive momentary reporting of pain with an electronic diary: reactivity, compliance, and patient satisfaction. *Pain*. 2003;104:343-351.
64. Mujagic Z, Leue C, Vork L, et al. The Experience Sampling Method–a new digital tool for momentary symptom assessment in IBS: an exploratory study. *Neurogastroenterol Motil*. 2015;27:1295-1302.
65. Weinland SR, Morris CB, Hu Y, Leserman J, Bangdiwala SI, Drossman DA. Characterization of episodes of irritable bowel syndrome using ecological momentary assessment. *Am J Gastroenterol*. 2011;106:1813-1820.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Smeets FGM, Masclee AAM, Conchillo JM, Keszthelyi D. Systematic review: Disease-specific instruments to assess gastrointestinal symptoms in functional dyspepsia. *Neurogastroenterol Motil*. 2018;30:e13327. https://doi.org/10.1111/nmo.13327