Neurogastroenterology & Motility

WILEY

Neurogastroenterology & Motility NGM

ORIGINAL ARTICLE

The Waiting Room Questionnaire, a Patient-Reported Outcome Instrument Including Pictograms for Facilitating the Diagnosis of Functional Gastrointestinal Disorders

Nick Goelen[1] | Janne Scheepers[1] | Florencia Carbone[1] | Lieselot Holvoet[2] | Alain Vandenberghe[3] | Joris Arts[2] | Guy Boeckxstaens[1] | Philip Caenepeel[2] | Ingrid Gijsen[4] | Pascale Latour[5] | Véronique Lefèbvre[6] | Joris Neirynck[7] | Hubert Piessevaux[8] | Tim Vanuytsel[1] | Philippe Vergauwe[9] | Jan Tack[1]

[1]Gastroenterology, TARGID, KU Leuven, Leuven, Belgium | [2]Gastroenterology, UZ Leuven, Leuven, Belgium | [3]Medical Research Consultant, Chaumont-Gistoux, Belgium | [4]Gastroenterology, Mariaziekenhuis Noord-Limburg, Pelt, Belgium | [5]Gastroenterology, CHU de Liège, Liège, Belgium | [6]Gastroenterology, CHR de Namur, Namur, Belgium | [7]Gastroenterology, Regional Hospital Heilig Hart, Leuven, Belgium | [8]Gastroenterology, Cliniques Universitaires St. Luc, Brussels, Belgium | [9]Gastroenterology, AZ Groeninge, Kortrijk, Belgium

Correspondence: Jan Tack (jan.tack@kuleuven.be)

Received: 19 October 2024 | **Revised:** 19 February 2025 | **Accepted:** 24 February 2025

Funding: This work was supported by the Methusalem grant from Leuven University to J.T. Mylan Belgium provided support for the organization of the investigator meetings for this project and the printing of the WRQ booklets. T.V. is a senior clinical research fellow of the Flanders Research Foundation (Research Foundation Flanders). N.G. was a SB PhD fellow at FWO (Research Foundation Flanders).

Keywords: epigastric pain syndrome | functional dyspepsia | meal-related symptoms | postprandial distress syndrome | Rome criteria

ABSTRACT

Introduction: Disorders of gut-brain interaction (DGBI) are a heterogeneous group of chronic conditions without a known organic etiology, with a significant socio-economic impact. In clinical practice, challenges remain in accurately assessing DGBI symptoms. Accurate interpretation and communication of the symptom pattern are imperative for a correct diagnosis. Hence, our group recently developed a Rome criteria-based Waiting Room Questionnaire (WRQ) with pictograms for visual support to improve symptom identification. This cross-sectional study aims to validate the WRQ for symptom assessment and DGBI diagnosis.

Methods: In a tertiary care multicentric study, 245 ambulatory patients (71% female, 41 ± 19.4 years old) completed the WRQ. The clarity and added value were recorded by the patients. The concordance between diagnosis based on the WRQ most bothersome symptom and the expert clinician's diagnosis were compared.

Results: Based on the expert opinion of clinicians, 53 patients were primarily diagnosed with gastro-esophageal reflux disease (GERD), 66 with postprandial distress syndrome (PDS), 36 with epigastric pain syndrome (EPS), 57 with irritable bowel syndrome (IBS) and 26 with chronic nausea/vomiting syndrome (CNVS). Pictograms entailed an added value according to 88% of the patients, and 95% of the patients understood all questions. Overall, patients' answers on the WRQ were highly reliable. Inter-rater agreement was substantial (Cohen's kappa = 0.72) with an overall concordance of 78% between expert opinion and the most bothersome symptom reported on the WRQ.

Conclusion: The WRQ is a comprehensive and reliable tool facilitating the diagnosis of patients with DGBI.

JT takes responsibility for the integrity of the work as a whole, from inception to published article. All authors approved the final version of the manuscript.

Summary

- Disorders of gut-brain interaction (DGBI) are challenging to diagnose, but accurate symptom interpretation is crucial for proper treatment.
- A new symptom assessment tool, the Waiting Room Questionnaire (WRQ), was developed to help patients describe their symptoms more clearly using pictograms.
- The WRQ was found to be reliable, with high agreement between patient responses and expert diagnoses, making it a reliable tool for facilitating the diagnosis of patients with DGBI.

1 | Introduction

Disorders of gut-brain interaction (DGBI) are highly prevalent condition, with individual diagnoses based on the Rome criteria [1]. Since no specific biomarkers exist, the diagnosis of the different DGBI is based on the patients symptom profiles. In the upper gastrointestinal (GI) tract, the main DGBI are functional dyspepsia (FD), with subgroups postprandial distress syndrome (PDS) and epigastric pain syndrome (EPS), functional dysphagia, rumination syndrome (RS) and cyclic vomiting syndrome (CVS) (prevalences ranging from 1.2% for CVS to 7.2% for FD) [1, 2]. Not without limitations, substantial overlap exists between symptom profiles, which are at least partially taken into account in the Rome diagnostic criteria [3]. However, management and treatment decisions usually depend on the identification of a primary DGBI diagnosis [4, 5], which is determined by an accurate assessment of the dominant symptomatic profile (frequency and severity of symptoms) as communicated by the patient to the specialist during history taking [4, 6]. Similarly, selection of patients for clinical trials also depends on the symptom pattern as determined by questionnaires with verbal descriptors.

However, patient-physician communication or identification of specific symptoms by a patient can be troublesome, as the medical concept of symptom-associated features is not always known or understood by patients. For instance, patients have difficulty understanding and distinguishing certain symptoms such as postprandial fullness, bloating, and early satiation. Our group has previously developed and validated a set of pictograms illustrating upper gastro-intestinal symptoms to supplement verbal descriptors used in questionnaires [5]. We documented that, in tertiary care FD patients, the addition of the pictograms to a Rome questionnaire significantly enhanced the agreement with the physician's interpretation of the symptom pattern compared to the same questionnaire without pictograms. Hence, pictograms have the potential for enhancing the comprehension, recall, and reporting of symptoms by the patient. Nevertheless, the gain in other settings than FD diagnosis and tertiary care needs further evaluation, and based on patients comments, some of the pictograms needed revision or redrawing [5].

The aim of the present study was to evaluate the benefit of and validate the "Waiting Room Questionnaire" (WRQ), a Rome criteria-based DGBI questionnaire with an updated set of pictograms. We studied the relevance, clarity and consistency of each of the question items and pictograms. In addition, we report the consistency and reliability of the WRQ to detect and differentiate DGBIs compared to in-depth interviews performed by expert clinicians.

2 | Materials and Methods

2.1 | Patient Selection and Study Design

For this cross-sectional observational study, consecutive ambulatory patients presenting with DGBI for an initial diagnostic consultation at one of 13 tertiary and secondary gastroenterology practices in Belgium (Dutch and French speaking) were eligible for the study. Patients were included if they reported one or several of the following digestive/abdominal symptoms: heartburn, nausea, vomiting, post-prandial fullness, early satiation, epigastric burning, epigastric pain, excessive belching, abdominal symptoms relieved by stool and/or gas movements, abdominal bloating, watery stools, hard stools, low and/or high frequency of stool movements, related to meals or not. Patients were included if no organic digestive or extra digestive, metabolic, biological, or iatrogenic disease was found to explain the symptoms. A negative endoscopic evaluation was required if necessary for the Rome diagnostic criteria.

Eligible patients completed the WRQ in the waiting room before their scheduled outpatient visit with the gastroenterologist. The questionnaire was made available in Dutch and French and was only provided to native Dutch- and French-speaking patients. After the outpatient visit, which included classical in-depth history taking and clinical examination, clinicians completed a separate case report form. The study protocol diagram is shown in Figure 1. Consecutive recruitment aimed to include at least 25 patients for each primary diagnostic group (irritable bowel syndrome (IBS), gastro-esophageal reflux disease (GERD), FD PDS, FD EPS, chronic nausea vomiting syndrome (CNVS)). GERD patients were also subdivided into patients on or off proton pump inhibitors (PPI). All study procedures were approved by the Ethics Committee of Leuven University Hospital, Belgium (reference number: S57433) and were performed in accordance with the Declaration of Helsinki. All patients signed an informed consent form before being included in the study.

2.2 | The Waiting Room Questionnaire

The WRQ contains 40 questions on DGBI-related GI symptoms, accompanied by the respective pictogram, updated from previous literature [5] illustrating the origin of the surveyed symptom (discomfort, pain, bloating, early satiation, postprandial fullness, nausea, heartburn and discomfort or pain relieved by a bowel movement) (Figure 2). In a screening population ($n = 10$), completing the WRQ took an average of 7 minutes.

FIGURE 1 | Study procedure diagram. Informed consent was obtained prior to a scheduled appointment with a gastroenterologist. Patients completed the Waiting room Questionnaire (WRQ) by themselves before the clinician's clinical in-depth history taking. Clinicians completed a separate case report form (CRF) after the consultation.

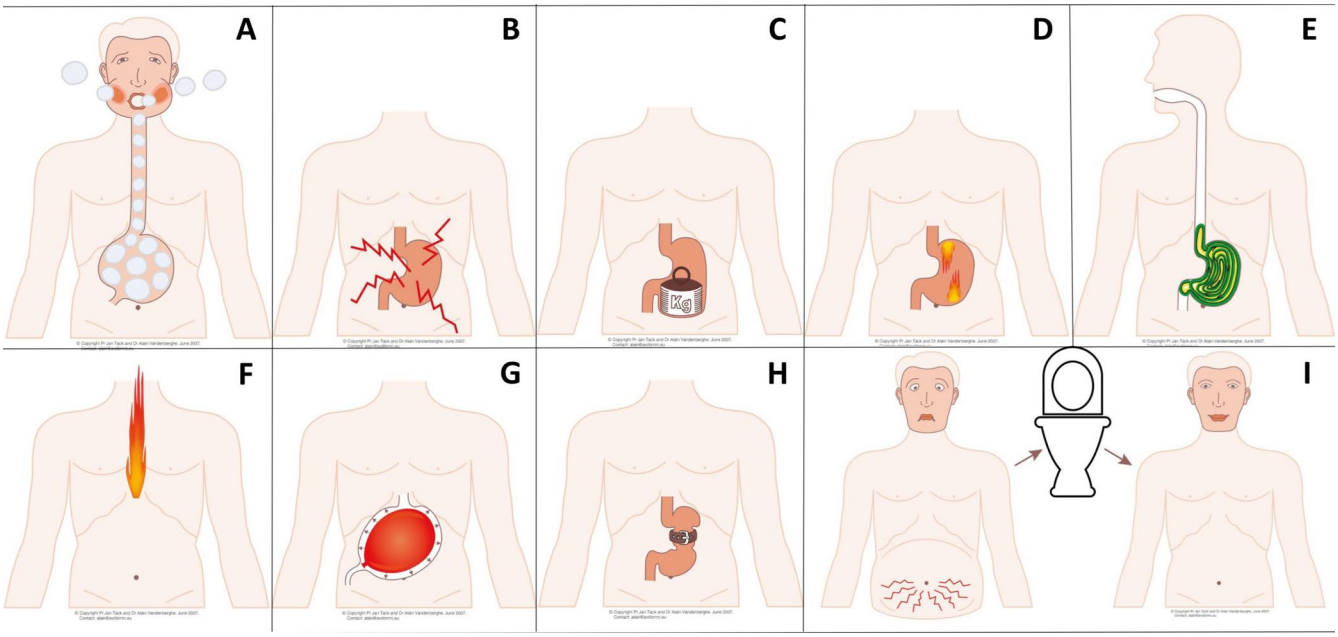


FIGURE 2 | Pictograms depicting (A) belching, (B) epigastric pain, (C) postprandial fullness, (D) epigastric burning, (E) nausea, (F) retrosternal burning, (G) bloating, (H) early satiation, (I) lower abdominal discomfort or pain relieved by defecation. Copyright Jan Tack and Alain Vandenberghe.

The WRQ is based on the Rome IV questionnaire for FD, CNVS, functional abdominal bloating, and belching disorders [1]. The Rome IV questionnaire includes questions related to:

- symptom presence (0: absent, 1: present),
- time of onset (0: less than 6 months ago, 1: more than 6 months ago),
- and their frequency (0–6; 0: never or less than once a month, 1: 1 day a month, 2: two to 3 days a month, 3: once a week, 4: two or 3 days a week, 5: most days, 6: every day).

For epigastric pain, in addition to the Rome criteria, severity was evaluated (on a scale of 1 to 5, with 1 indicating very mild and 5 signifying very severe discomfort), as they represent distinct aspects regarding epigastric pain [7].

Furthermore, we considered the relationship of epigastric pain, nausea, and belching to meals (on a scale of 0 to 5, ranging from never (0) to always [5]).

For IBS, the Rome III questionnaire was used. The decision to use Rome IV for upper GI disorders and Rome III for IBS was

driven by the availability of the updated upper GI criteria to the researchers at the time of study design, with only minimal differences between Rome III and IV for upper GI disorders.

To enable the diagnosis of functional heartburn (FH)/gastroesophageal reflux disease (GERD), a previously validated questionnaire was incorporated into the WRQ [8]. A positive answer on all questions is a reliable predictor of pathological pH-monitoring (76% specificity) [8].

In addition, patients were asked to identify their most bothersome symptom.

Moreover, the patient was asked about weight loss, concomitant medication, and the presence of other gastrointestinal symptoms that were not addressed in the questions.

2.3 | Cognitive Evaluation of the Waiting Room Questionnaire

To evaluate the relevance, clarity, and consistency of each of the question items and pictograms, a cognitive evaluation was conducted using a paper questionnaire, which was included as the last page of the WRQ form. Patients were asked

1. Whether they understood all questions clearly and
2. Whether the pictograms aided in the understanding of the questions.

In case patients replied "No", further clarification was requested per the misunderstood question. Patients indicated the presence of any additional abdominal symptoms that were not addressed.

2.4 | Evaluation of the Diagnostic Value of the Waiting Room Questionnaire

After history and clinical examination, the clinician, who was blinded to the patient's WRQ answers, indicated the primary diagnoses on the case report form (CRF). The physician was also allowed to add additional secondary overlapping diagnoses if he/she felt this was relevant to the clinical assessment. Once this was noted, the clinicians were unblinded to the WRQ responses of the patients and evaluated to what degree the patient's written answers were reliable based on the history taking. If the patient indicated additional or overlapping symptoms that had not been addressed by the clinician, this could be shortly evaluated for confirmation by additional questions from the clinician. Reliability was scored in four different categories:

1. overall symptoms (WRQ, symptom severity, FH/GERD questionnaire),
2. meal-relationship of epigastric pain, nausea and belching
3. stool pattern as addressed in the Rome III IBS criteria,
4. and the most bothersome symptom.

The level of reliability was indicated on scales from "Definitely" (>90% of the items addressed on the WRQ were confirmed during the interview) to "Not at all" (<20% of the items addressed on the WRQ were confirmed during the interview). At the end of the interaction, clinicians registered a primary diagnosis and potential secondary diagnoses. To register diagnoses, the clinicians could choose one DGBI as primary diagnosis but were able to tick none to multiple DGBIs as secondary diagnoses.

2.5 | Data Analysis and Statistics

Three diagnoses were derived for analysis: (i) The clinician's primary diagnosis after in-depth history taking, hence based on clinical expertise alone. This was considered the *clinical reference diagnosis*. (ii) The *predominant diagnosis* based on the patient's most bothersome symptom, as surveyed in the WRQ. (iii) The diagnosi(e)s based on the responses of the patient to the WRQ, taking into account the diagnostic thresholds and guidelines defined in the Rome III or IV criteria for upper GI disorders or IBS, respectively [1, 9]. The heartburn questionnaire was considered positive if all four questions were answered positively.

Multiple DGBI diagnoses could be present and confirmed in a single patient. For the descriptive analysis, patients were grouped by their clinical reference diagnosis.

Inter-rater agreement between the clinical reference diagnosis and the most bothersome symptom identified by the patient on the WRQ was calculated using Cohen's Kappa ($k=(p_o-p_e)/(1-p_e)$ with p_o = relative observed agreement among raters; p_e = hypothetical probability of chance agreement). Data are represented as mean (SD).

Additionally, as there is known overlap between EPS and PDS, the concordance of FD reported by patients and as clinical reference diagnosis was evaluated.

The use of PPI can also have an impact on patient's symptom reporting, especially for FH and GERD. Therefore, the concordance of diagnosis reported by patients and as clinical reference diagnosis was studied in subgroups of patients on PPI and off PPI.

Results are shown as mean ± SD where applicable. Statistical tests were done using Matlab R2016 and Graphpad.

3 | Results

3.1 | Demographics

A total of 258 ambulatory patients were recruited. After the omission of incomplete questionnaires ($n=13$), 245 were suitable for analysis (71% women, average age 41 ± 19.4 years).

After history taking and clinical evaluation, clinicians diagnosed 53 patients with predominant FH/GERD, 66 with PDS, 36 with EPS, 57 with IBS, and 26 with CNVS (Table 1). The clinician diagnosis was missing for three patients, and four patients were diagnosed with a belching disorder. Of the 245 patients,

49% were taking a PPI or H_2-receptor antagonist. Patients with IBS took significantly less acid-suppressing medication compared to the FH/GERD subgroup ($p = 0.007$) and the PDS subgroup ($p = 0.03$). In total, 30% of the patients experienced weight loss (average: 7 ± 5.5 kg) since the onset of their symptoms (Table 1). Patients diagnosed predominantly with CNVS, reported weight loss significantly more often than patients with FD/GERD ($p = 0.03$).

Of the 53 patients diagnosed predominantly with FH/GERD (60.4% female, 41 ± 22.7 years old), 32 (60.4%) were on acid-suppressive therapy (PPI or H2-receptor antagonist, H2RA) (Table 2).

3.2 | Agreement Between the Clinicians' Primary Diagnosis and the WRQ

Primary diagnosis by the physician was defined in 241 patients (Table 3). Two patients were considered to have other diagnoses and for one patient, the physician could not select one specific diagnosis. The most bothersome symptom reported in the WRQ was available for 231 patients.

According to the responses of the patient to the WRQ, taking into account the diagnostic thresholds and guidelines defined in the Rome criteria, 102 patients were diagnosed with FH/GERD, 94 with PDS,

TABLE 1 | Demographics and clinical reference diagnosis.

Demographics		Clinical reference diagnosis				
	Total	FH/GERD	PDS	EPS	IBS	CNVS
N	238	53	66	36	57	26
Female (%)	71% (*n* = 169)	60.4% (*n* = 32)	72.7% (*n* = 48)	80.6% (*n* = 29)	78.9% (*n* = 45)	57.7% (*n* = 15)
Age	41 (± 19.4)	41 (± 22.7)	44 (± 16.2)	43 (± 18.5)	37 (± 21.8)	40 (± 13.4)
Usage of PPI or H_2-receptor antagonist (%)	48.7% (*n* = 116)	60.4% (*n* = 32)	53.0% (*n* = 36)	47.2% (*n* = 17)	33.3% (*n* = 19)	53.8% (*n* = 14)
Weight loss (%)	30% (*n* = 72)	18.9% (*n* = 10)	33.3% (*n* = 22)	36.1% (*n* = 13)	28.1% (*n* = 16)	42.3% (*n* = 11)

TABLE 2 | Demographics of FH/GERD patients.

	Total	On PPI/H2RA	Off PPI/H2RA
N	53	32	21
Female (%)	60.4% (*n* = 32)	62.5% (*n* = 20)	61.9% (*n* = 12)
Age (mean ± SD)	41 (± 22.7)	40 (± 26.9)	44 (± 14.6)
Weight loss (%)	18.9% (*n* = 10)	18.8% (*n* = 6)	19.0% (*n* = 4)

TABLE 3 | Patient categorization according to different diagnostic methods.

	Diagnostic method			% Agreement	
Diagnosis (*n*)	Clinician primary diagnosis	Most bothersome symptom in WRQ	WRQ with pictograms	Primary diagnosis and WRQ with pictograms	Primary diagnosis and most bothersome symptom
GERD/FH	53	42	102	69.4%	92.0%
PDS	66	57	94[a]	64.9%	85.7%
EPS	36	46	67[a]	74.3%	81.2%
IBS	57	49	32	77.6%	93.3%
CNVS	26	37	67	75.1%	94.5%
No diagnosis	0	0	56		
Total	238	231	362	**69.9%**[b]	**77.7%**[b]
Missing	4	14	0		

[a]EPS-PDS overlap in 44 patients.
[b]$P > 0.05$.

67 with EPS, 32 with IBS, and 67 with CNVS. A total of 56 patients did not qualify as having a DGBI according to the Rome criteria. As multiple diagnoses are possible, 101 patients showed overlap between one or more DGBI, and for 88 patients, one specific DGBI could be identified as per Rome IV diagnosis criteria. The largest overlap was found between GERD/FH and PDS (45 patients), PDS and EPS (44 patients) and GERD/FH and EPS (30 patients).

Agreement between the primary diagnosis made by the clinicians, the diagnosis obtained by analyzing the WRQ as per Rome IV criteria, and the reported most bothersome symptom in the WRQ is visualized in Figure 3. The best agreement is found between the clinicians' primary diagnosis and the most bothersome symptom, as reported in the WRQ.

The proportion of symptoms in each diagnostic group, based on the physician's primary diagnosis, is shown in Figure 4. The presence of each symptom was determined based on the Rome IV/III diagnostic threshold values for symptom frequency, as assessed in the WRQ. Lower abdominal pain/discomfort, postprandial fullness, and epigastric pain are the most specific indicators for the respective diagnosis of IBS, PDS, and EPS.

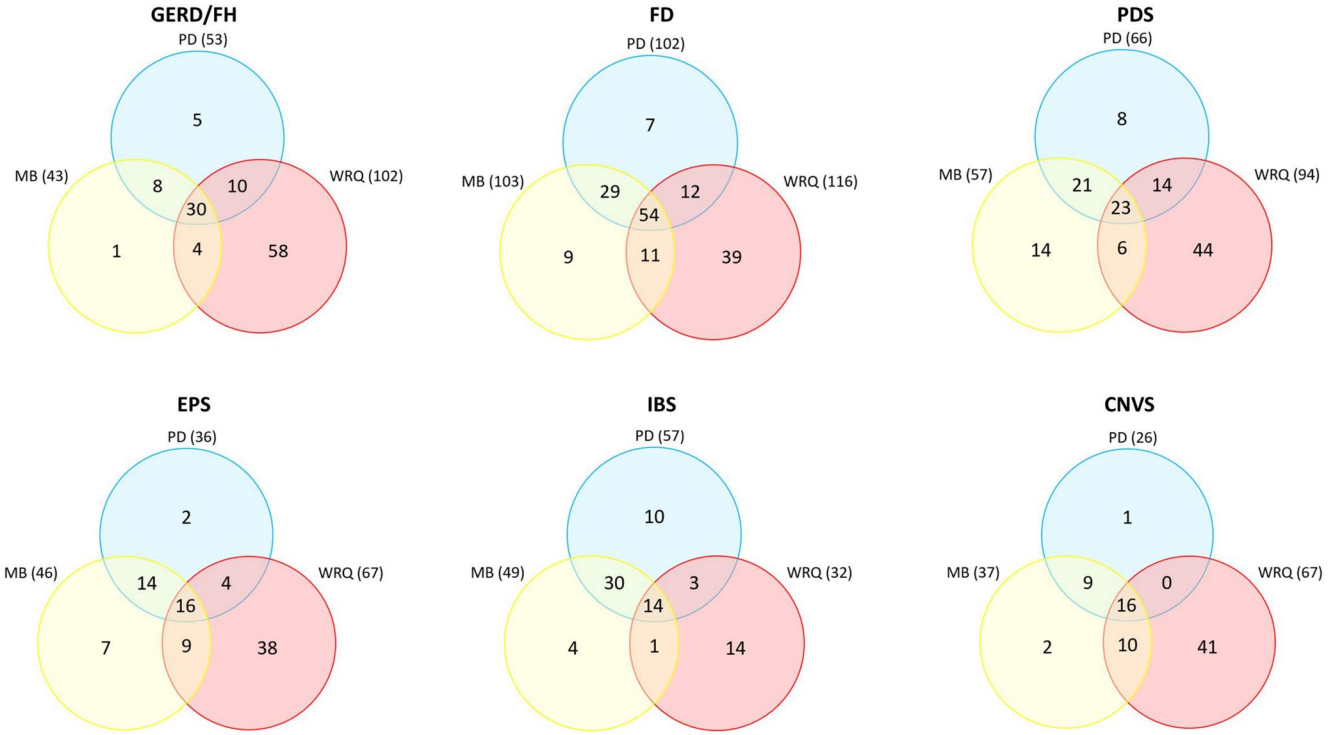
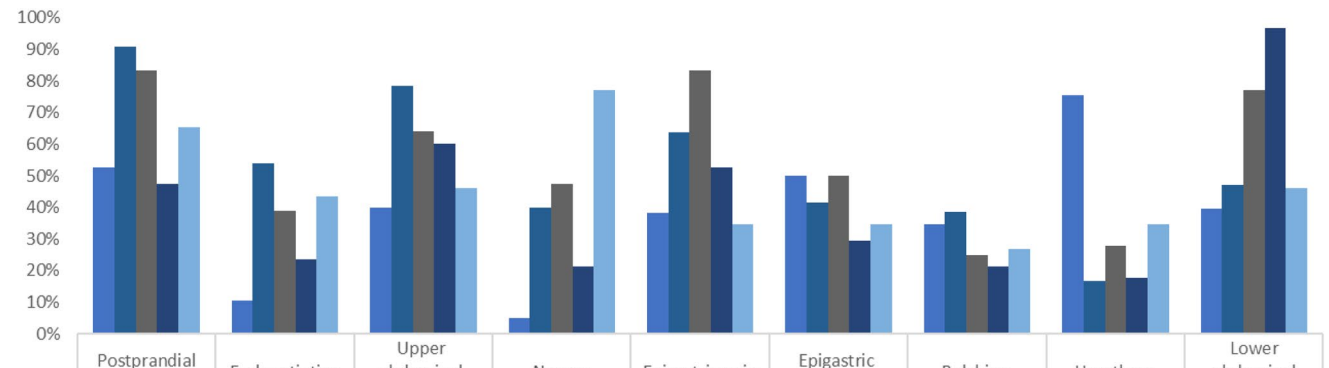


FIGURE 3 | Agreement of diagnostic methods. Numbers depict the absolute number of patients receiving a diagnosis as the clinicians' Primary diagnosis (PD), Most bothersome symptom (MB) and/or the full symptom assessment in the WRQ.



	Postprandial fullness	Early satiation	Upper abdominal bloating	Nausea	Epigastric pain	Epigastric burning	Belching	Heartburn	Lower abdominal pain/disconfort
GERD	53%	11%	40%	5%	38%	50%	35%	75%	40%
PDS	91%	54%	78%	40%	64%	42%	38%	17%	47%
EPS	83%	39%	64%	47%	83%	50%	25%	28%	77%
IBS	47%	24%	60%	21%	53%	29%	21%	18%	96%
CNVS	65%	43%	46%	77%	35%	35%	27%	35%	46%

FIGURE 4 | Prevalence of symptoms exceeding Rome IV thresholds. Bars represent the fraction of patients in each clinical reference diagnostic group with symptom scores exceeding the Rome IV/III diagnostic thresholds (i.e., frequent symptoms).

Table S1 presents the sensitivity and specificity of the WRQ as defined by the Rome criteria, together with the sensitivity and specificity of the most bothersome symptom as a diagnostic tool, using the clinicians' primary diagnosis as the reference standard. The highest sensitivity was noted for EPS according to the WRQ (83.3%, 95% CI: 67.19%–93.63%), whilst IBS according to the WRQ (29.8%, 95% CI: 18.43%–43.40%) was the least sensitive. The highest specificity was noted for GERD/FH according to the most bothersome symptom (97.4%, 95% CI: 94.03%–99.15%). FD according to the WRQ (65%, 95% CI: 56.62%–72.81%) ranked lowest in terms of specificity.

3.2.1 | Concordance Between Primary Clinical Diagnosis and the Most Bothersome Symptom Reported in the WRQ

The agreement between the clinical reference diagnosis and the most bothersome symptom identified by the patient on the WRQ was assessed for inter-rater consistency (Figure 5). A substantial agreement was observed between the clinical reference diagnosis and the patients' predominant diagnosis, based on the most bothersome symptom as surveyed in the WRQ. The number of observed agreements was 185 (77.7% of the observations, κ: 0.72 (95% CI: 0.66–0.79)).

Concordance was highest for CNVS (94.5% of the observations, κ: 0.76, 95% CI: 0.64–0.89) and IBS (93.3% of the observations, κ: 0.8, 95% CI: 0.71–0.90), followed by FH/GERD (92.0% of the observations, κ: 0.75, 95% CI: 0.65–0.86), EPS (91.2% of the observations, κ: 0.69, 95% CI: 0.57–0.81) and PDS (85.7% of the observations, κ: 0.63, 95% CI: 0.51–0.74) (Figure 5).

The inter-rater agreement for FD as one cohort was substantial, with 84.1% observed agreements between the clinical reference diagnosis and the most bothersome symptom (κ: 0.67, 95% CI: 0.58–0.77).

3.2.2 | Concordance Between the Primary Clinical Diagnosis and the Most Bothersome Symptom in the WRQ for Patients on and off Acid Suppressive Therapy

Inter-rater agreements were calculated between the clinical reference diagnosis and the patient's identified most bothersome symptom within subgroups categorized based on acid-suppressive therapy status. Substantial agreements were observed in both the subgroup receiving acid-suppressive therapy (73% agreement, κ: 0.66, 95% CI: 0.56–0.76) and the subgroup without acid-suppressive therapy (77% agreement, κ: 0.71, 95% CI: 0.62–0.80).

Remarkably, among patients with FH/GERD as their primary diagnosis, there was an almost perfect inter-rater agreement (96.8% agreement, κ: 0.876, 95% CI: 0.76–0.995) between the clinical reference diagnosis and the patient-identified most bothersome symptom in the subgroup without acid-suppressive therapy. Among patients on acid-suppressive therapy, a substantial level of agreement persisted (88.4% agreement, κ: 0.64, 95% CI: 0.47–0.80).

3.2.3 | Concordance Between the Primary Clinical Diagnosis and the WRQ Diagnosis as Defined by Rome Criteria

Inter-rater agreement between the clinical reference diagnoses and the WRQ diagnoses (based on the Rome IV criteria and pictograms) was calculated. The highest concordance was observed for IBS (77.6% agreement, $\kappa = 0.56$, 95% CI: 0.12–0.4) and CNVS (75.1% agreement, $\kappa = 0.23$, 95% CI = 0.09–0.35), followed

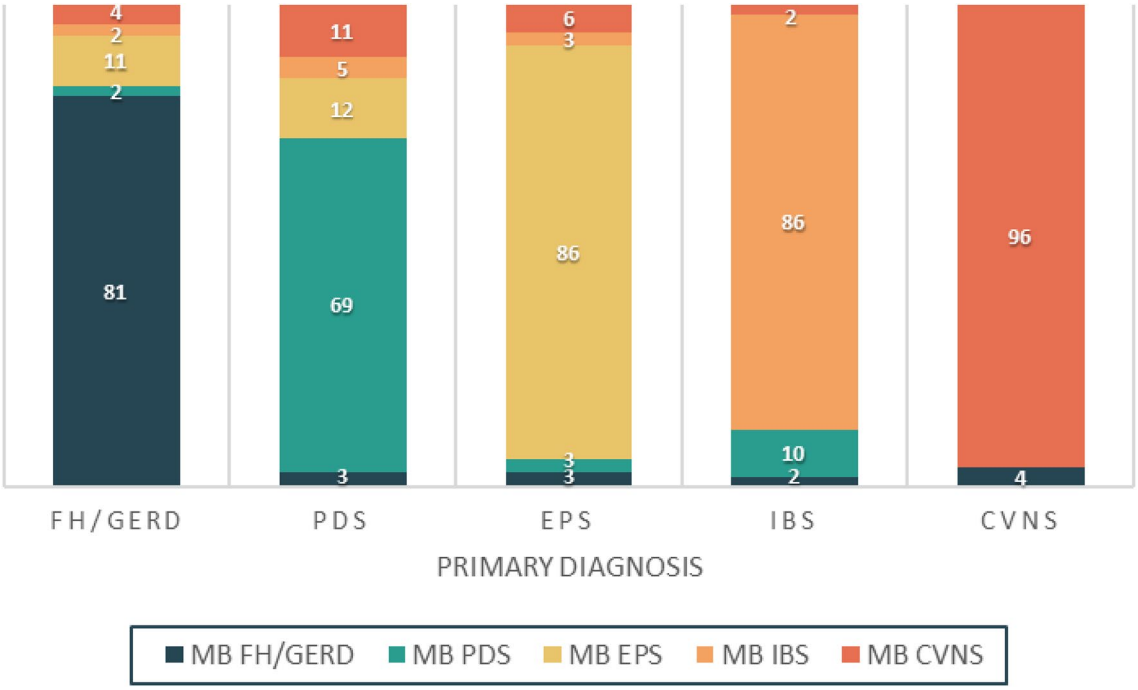


FIGURE 5 | Inter-rater agreement between the primary diagnosis and the diagnosis based on the most bothersome symptom. All patients are grouped in bars according to the primary diagnosis (based on expert opinion). Each bar is subdivided according to the diagnosis based on their most bothersome symptom.

by EPS (74.3% agreement, $\kappa = 0.24$, 95% CI: 0.11–0.37), FH/GERD (69.4% agreement, $\kappa = 0.32$, 95% CI: 0.21–0.43) and PDS (64.9% agreement, $\kappa = 0.21$, 95% CI: 0.09–0.34).

With both the PDS and EPS subgroups polled together, the physicians diagnosed 102 patients with FD. An inter-rater agreement of 64.9% was observed (159 agreements, $\kappa = 0.29$, 95% CI: 0.17–0.41) between the clinical reference diagnosis of FD and the WRQ.

3.2.4 | Concordance Between the Clinical Reference Diagnosis and the WRQ Diagnosis for Patients on and off Acid Suppressive Therapy

For FH/GERD, inter-rater agreements were calculated between the clinical reference diagnosis and the diagnosis from the WRQ within subgroups categorized based on acid-suppressive therapy status. Among patients diagnosed primarily with FH/GERD there was 79.4% inter-rater agreement (100 agreements, κ: 0.46, 95% CI: 0.30–0.63) between the clinical reference diagnosis and the WRQ in the subgroup without acid-suppressive therapy. Among patients on acid-suppressive therapy, a slight agreement was found, with 58.8% agreement (70 agreements, κ: 0.19, 95% CI: 0.03–0.34).

3.2.5 | Determination of Additional Diagnoses by the WRQ and Their Concordance

Accounting for secondary diagnoses, clinicians assigned a total of 504 diagnoses. Among these, 260 cases (52%) were confirmed by the waiting room questionnaire. When calculating the inter-rater agreement between the clinical reference diagnosis and the WRQ diagnosis, taking the additional diagnoses into account, the highest concordance was found for FH/GERD with 77.6% (190 agreements, κ: 0.54, 95% CI: 0.54–0.05). This was followed by CNVS with 75.9% agreement (186 agreements, κ: 0.31, 95% CI: 0.17–0.44), PDS with 66.5% agreement (163 agreements, κ: 0.34, 95% CI: 0.23–0.45), EPS with 64.5% agreement (158 agreements, κ: 0.22, 95% CI: 0.09–0.34) and IBS with 60.4% agreement (148 agreements, κ: 0.20, 95% CI: 0.12–0.29).

3.2.6 | Content Validity and Cognitive Evaluation of the WRQ

In total, 231 patients responded to the cognitive assessment regarding the question clarity; the question on the added value of the pictograms was completed by 224 patients. The pictograms entailed an added value according to 87.5% of the responders, and 85.2% of responders indicated they understood all questions. The highest reporting by patients of unclarity of a pictogram was for the nausea pictogram, but still only in a small minority of the subjects ($n = 4$). Clinicians indicated that > 90% of the written answers were confirmed by their interview in 77.1% of the cases. Patients were considered highly reliable in 86.1% of the cases with regard to meal-related symptoms and in 86.9% with regard to bowel movements according to the physician's interview (Table S2).

4 | Discussion

The present study investigated the diagnostic agreement between clinicians' assessments and patient-reported symptoms using a waiting room questionnaire (WRQ) in the context of ambulatory gastroenterology clinic patients.

The WRQ was completed by 245 ambulatory patients while waiting for their consultation with an expert gastroenterologist. The WRQ was well received by the patients as well as the clinicians. Pictograms entailed an added value for 87.5% of the patients who responded, and 85.2% of the patients who responded understood all questions. Furthermore, clinicians indicated that the vast majority of the written answers in the WRQ were confirmed during their interview, resulting in a good overall concordance for the predominant symptomatic profile identified by the WRQ and by the clinician (77.1%). The current data allowed a more comprehensive characterization of the symptomatic profiles, resulting in a higher percent agreement and kappa value compared to our previous pictogram study (55%, $\kappa = 0.622$) [5].

The concordance between the clinical reference diagnosis and the symptom identified as most bothersome by the patient on the WRQ was evaluated. A substantial agreement was noted between the clinical reference diagnosis and the predominant symptom reported by patients on the WRQ. Concordance was highest for CNVS, with a 94.5% agreement. For the entire functional dyspepsia population, the agreement was substantial at 84.1%, but when analyzing the subgroups, the agreement was higher for EPS (91.2%) than for PDS (85.7%). Based on the Rome IV criteria, 18% of patients exhibited overlap between PDS and EPS. GERD/FH showed a substantial agreement of 92.0% between the primary diagnosis and the most bothersome symptom.

The high concordance between the most bothersome symptom and clinical reference diagnosis for IBS (93.3%) can, in part, be attributed to the distinct anatomical localization of the associated symptoms. In contrast, only 53% of the patients with a primary diagnosis of IBS did meet the Rome III criteria. In the updated Rome IV criteria, the thresholds for IBS are even more stringent [10]. This could potentially reduce false-positive IBS diagnoses in the questionnaire, thereby slightly improving inter-rater agreement.

The use of PPI therapy was high in the GERD patients (60.4%) compared to the other groups (33.3%–53.8%). Notably, for patients diagnosed primarily with FH/GERD, an almost perfect inter-rater agreement was observed among patients off acid-suppressive therapy. Altered symptom patterns and symptom relief caused by acid-suppressive therapy could partially explain the differences in inter-rater agreements.

Based on the good inter-rater agreement between the primary diagnosis of the physician and the most bothersome symptom reported by the patients, one might argue that the most bothersome symptom alone is a suitable indicator to guide the diagnosis of these frequent DGBI. This highlights the potential value of having patients select their most bothersome symptom, a step often not implemented in routine clinical practice. It should be

noted that patients were able to reflect on their symptoms when answering the earlier questions in the WRQ and were provided with pictograms to help clarify the symptoms, aiding them in identifying the symptom that troubled them most. However, our data also clearly indicate that the in-depth history taking by a trained physician enables a more holistic comprehension of a patient's individual symptom pattern and treatment history, resulting in a more refined final diagnosis.

The WRQ demonstrated higher sensitivity for identifying both FD and IBS compared to the Rome IV and Rome III questionnaires, respectively. For FD, the sensitivity of the WRQ using the most bothersome symptom was 81.4%, significantly outperforming the Rome IV questionnaire at 54.7% [11]. However, this improved sensitivity came with a reduction in specificity (86.0% vs. 93.3%). For IBS, the WRQ showed a slightly higher sensitivity (77.2% vs. 68.8%) and higher specificity (97.3% vs. 79.5%) than the Rome III questionnaire [12]. It should be noted that the Rome III/IV questionnaires were validated in a larger sample within a secondary care setting.

The WRQ provides key benefits through its use of pictograms, which enhance patient understanding and reporting of symptoms, as previously reported [5]. Additionally, the WRQ combines the diagnostic questions from Rome questionnaires with measures of symptom severity, meal-relatedness, and a validated FH/GERD questionnaire. This design allows for a more complete assessment of gastrointestinal symptoms, helping to identify overlapping or complex conditions that might otherwise go undiagnosed. The WRQ is potentially of great value in supporting the diagnosis of DGBIs in more complex patients in clinical practice, as well as screening the suitability of symptoms when considering the recruitment of patients for DGBI therapeutic trials. The instrument does not provide a severity rating of the respective diagnosis, as this requires other diagnosis-specific patient-reported outcome questionnaires.

A limitation of the current study is its cross-sectional nature, but this is inherent to the setting at the diagnostic stage. We did not assess test–retest variability and changes in diagnostic features during treatment, but these are beyond the scope of assessing performance in an outpatient diagnostic setting. Although physicians were blinded to the WRQ results, their familiarity with the questionnaire's structure may have influenced their clinical assessments. This potential Hawthorne effect could have led to a higher concordance by prompting physicians to focus on symptoms emphasized in the WRQ.

Another limitation of this study is that patient recruitment took place in tertiary centers, which may limit the generalizability of the findings. Additionally, the study includes relatively small sample sizes for each DGBI.

Taken together, the performance of this self-administered patient questionnaire in a secondary or tertiary care setting is attractive. Given its reliability in capturing predominant symptom profiles, the WRQ could aid in history-taking during clinical consultations. Moreover, it has potential as a screening tool for clinical trials targeting DGBIs. We are currently in the process of conducting a follow-up study in primary care patients to assess the added value of the WRQ for general practitioners.

5 | Conclusion

We conclude that the waiting room questionnaire is a comprehensive and easy to use tool that facilitates the screening and diagnosing of frequent DGBIs in gastroenterology practice. Implementation in daily clinical practice, as well as in clinical research, might result in a more standardized and time-saving process.

Author Contributions

A.V., F.C., L.H., and J.T.: study concept and design. A.V., F.C., J.T., L.H., and N.G.: data acquisition; F.C., J.T., L.H., N.G., and J.S.: data analysis and interpretation. J.A., G.B., P.C., I.G., P.L., V.L., J.N., H.P., T.V., P.V., and J.T.: expert clinicians that participated in the trial. N.G., F.C., J.S., and J.T.: drafting of the manuscript. All authors: critical revision of the manuscript.

Acknowledgments

The authors thank Dr. Dirk Staessen, Dr. Ignace Vanmoerkerke and Dr. Steven Van Outryve for their contributions to this study. Funding was provided by a Methusalem grant from Leuven University to J.T. Mylan Belgium provided support for the organization of the investigator meetings for this project and the printing of the WRQ booklets. T.V. is a senior clinical research fellow of the Flanders Research Foundation (Research Foundation Flanders). N.G. was a SB PhD fellow at FWO (Research Foundation Flanders).

Conflicts of Interest

The authors declare no conflicts of interest.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author upon reasonable request.

References

1. "Rome IV Diagnostic Criteria for FGIDs," https://theromefoundation.org/rome-iv/rome-iv-criteria/.

2. A. D. Sperber, S. I. Bangdiwala, D. A. Drossman, et al., "Worldwide Prevalence and Burden of Functional Gastrointestinal Disorders, Results of Rome Foundation Global Study," *Gastroenterology* 160, no. 1 (2021): 99–114.

3. N. J. Talley, M. M. Walker, and G. Holtmann, "Functional Dyspepsia," *Current Opinion in Gastroenterology* 32, no. 6 (2016): 467–473.

4. J. Tack, R. Bisschops, and G. Sarnelli, "Pathophysiology and Treatment of Functional Dyspepsia," *Gastroenterology* 127 (2004): 1239–1255.

5. J. Tack, F. Carbone, L. Holvoet, H. Vanheel, T. Vanuytsel, and A. Vandenberghe, "The Use of Pictograms Improves Symptom Evaluation by Patients With Functional Dyspepsia," *Alimentary Pharmacology & Therapeutics* 40, no. 5 (2014): 523–530.

6. J. Tack, N. J. Talley, M. Camilleri, et al., "Functional Gastroduodenal Disorders," *Gastroenterology* 130, no. 5 (2006): 1466–1479.

7. F. Carbone, L. Holvoet, T. Vanuytsel, and J. Tack, "Rome III Functional Dyspepsia Symptoms Classification: Severity vs Frequency," *Neurogastroenterology and Motility* 29, no. 6 (2017): 13024.

8. J. Tack, P. Caenepeel, J. Arts, K.-J. Lee, D. Sifrim, and J. Janssens, "Prevalence of Acid Reflux in Functional Dyspepsia and Its Association With Symptom Profile," *Gut* 54, no. 10 (2005): 1370–1376.

9. D. A. Drossman, "Rome III Diagnostic Criteria for Functional Gastrointestinal Disorders," 2006.

10. C. J. Black, Y. Yiannakou, L. A. Houghton, and A. C. Ford, "Epidemiological, Clinical, and Psychological Characteristics of Individuals With Self-Reported Irritable Bowel Syndrome Based on the Rome IV vs Rome III Criteria," *Clinical Gastroenterology and Hepatology* 18, no. 2 (2020): 392–398.

11. O. S. Palsson, W. E. Whitehead, M. A. L. Van Tilburg, et al., "Development and Validation of the Rome IV Diagnostic Questionnaire for Adults," *Gastroenterology* 150, no. 6 (2016): 1481–1491.

12. A. C. Ford, P. Bercik, D. G. Morgan, C. Bolino, M. I. Pintos–Sanchez, and P. Moayyedi, "Validation of the Rome III Criteria for the Diagnosis of Irritable Bowel Syndrome in Secondary Care," *Gastroenterology* 145, no. 6 (2013): 1262–1270.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.