















# Standardized system and App for continuous patient symptom logging in gastroduodenal disorders: Design, implementation, and validation

Gabrielle Sebaratnam<sup>1,2</sup>  | Nikita Karulkar<sup>1</sup>  | Stefan Calder<sup>1,2</sup>  |  
Jonathan S.T. Woodhead<sup>2,3</sup>  | Celia Keane<sup>1,2</sup>  | Daniel A. Carson<sup>1</sup>  | Chris Varghese<sup>1</sup>  |  
Peng Du<sup>1,2</sup>  | Stephen J. Waite<sup>2</sup>  | Jan Tack<sup>4</sup>  | Christopher N. Andrews<sup>2,5</sup>  |  
Elizabeth Broadbent<sup>1</sup>  | Armen A. Gharibans<sup>1,2</sup>  | Greg O'Grady<sup>1,2</sup> 

<sup>1</sup>The University of Auckland, Auckland, New Zealand

<sup>2</sup>Alimetry Ltd, Auckland, New Zealand

<sup>3</sup>Maurice Wilkins Centre for Molecular Biodiscovery, Auckland, New Zealand

<sup>4</sup>Department of Gastroenterology, University Hospitals, Leuven, Belgium

<sup>5</sup>Division of Gastroenterology, Cumming School of Medicine, University of Calgary, Calgary, Alberta, Canada

## Correspondence

Armen A. Gharibans, Auckland Bioengineering Institute and Department of Surgery, University of Auckland, Auckland, New Zealand.

Email: armen.gharibans@auckland.ac.nz

Greg O'Grady, Department of Surgery, University of Auckland, Auckland, New Zealand.

Email: greg.ogrady@auckland.ac.nz

## Funding information

John Mitchell Crouch Fellowship from the Royal Australasian College of Surgeons; Health Research Council of New Zealand

## Abstract

**Background:** Functional gastroduodenal disorders include functional dyspepsia, chronic nausea and vomiting syndromes, and gastroparesis. These disorders are common, but their overlapping symptomatology poses challenges to diagnosis, research, and therapy. This study aimed to introduce and validate a standardized patient symptom-logging system and App to aid in the accurate reporting of gastroduodenal symptoms for clinical and research applications.

**Methods:** The system was implemented in an iOS App including pictographic symptom illustrations, and two validation studies were conducted. To assess convergent and concurrent validity, a diverse cohort with chronic gastroduodenal symptoms undertook App-based symptom logging for 4 h after a test meal. Individual and total post-prandial symptom scores were averaged and correlated against two previously validated instruments: PAGI-SYM (for convergent validity) and PAGI-QOL (for concurrent validity). To assess face and content validity, semi-structured qualitative interviews were conducted with patients.

**Key Results:** App-based symptom reporting demonstrated robust convergent validity with PAGI-SYM measures of nausea ( $r_s = 0.68$ ), early satiation ( $r_s = 0.55$ ), bloating ( $r_s = 0.48$ ), heartburn ( $r_s = 0.47$ ), upper gut pain ( $r_s = 0.40$ ), and excessive fullness ( $r_s = 0.40$ ); all  $p < 0.001$  ( $n = 79$ ). The total App-reported Gastric Symptom Burden Score correlated positively with PAGI-SYM ( $r_s = 0.56$ ; convergent validity;  $p < 0.001$ ), and negatively with PAGI-QOL ( $r_s = -0.34$ ; concurrent validity;  $p = 0.002$ ). Interviews demonstrated that the pictograms had adequate face and content validity.

**Conclusions and Inferences:** The continuous patient symptom-logging App demonstrated robust convergent, concurrent, face, and content validity when used within a 4-h post-prandial test protocol. The App will enable standardized symptom reporting and is anticipated to provide utility in both research and clinical practice.

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2022 The Authors. Neurogastroenterology & Motility published by John Wiley & Sons Ltd.

**KEYWORDS**

chronic nausea and vomiting syndrome, digital health, functional disorders, functional gastrointestinal disorders, gastroparesis, gut-brain-axis

**1 | INTRODUCTION**

Functional dyspepsia (FD) and chronic nausea and vomiting syndromes (CNVS) affect 7.2% and 1.2% of the global population, respectively, and significantly impact quality of life.<sup>1-3</sup> According to the Rome IV Criteria, FD is characterized by excessive fullness and early satiation (dominant in the post-prandial distress syndrome subtype), and epigastric pain and/or burning (dominant in the epigastric pain syndrome subtype), while CNVS patients predominantly experience nausea and vomiting.<sup>4</sup> However, these syndromes and symptoms often co-exist, while also overlapping with gastroparesis, which is controversially distinguished by the presence of delayed gastric emptying.<sup>5,6</sup> Additional symptoms such as bloating and belching are commonly present in affected patients,<sup>4</sup> while FD frequently also co-exists with gastro-oesophageal reflux disease (GERD).<sup>7</sup>

Distinguishing these disorders remains challenging owing to these overlaps and the ongoing lack of objective and specific biomarkers. A clear clinical characterization of specific symptoms is therefore essential for diagnosis, together with the exclusion of organic pathologies. A distinction must also be made between other potentially co-existing functional / gut-brain-axis disorders such as irritable bowel syndrome.<sup>8</sup> Moreover, accurate characterization is subject to the quality of clinical communication and may be negatively impacted by use of jargon, constraints on clinical time, and inaccuracy in patient recall of their symptom experiences.<sup>9</sup> Pictograms have been shown to assist in the understanding and communication of gastric symptoms between patients and healthcare providers, improving symptom reporting accuracy.<sup>9,10</sup> However further validation of pictogram use is desirable before they can be reliably integrated into clinical practice.

Validated instruments based on symptom recall are already available for longer-term assessments and are commonly used in research contexts, such as the Patient Assessment of Upper Gastrointestinal Symptom Severity Index (PAGI-SYM), Gastroparesis Cardinal Symptom Index (GCSI), and GCSI Daily Diary.<sup>11-13</sup> However, a continuous reporting tool that enables construction of a real-time symptom profile is also desirable to allow comparison with concurrent diagnostic tests, investigate provocations, and evaluate interventions. Continuous granular symptom profiling is also particularly important in functional gastrointestinal (GI) disorders because temporal correlations comprise part of the "Plausibility Criteria" that are recommended for use in the evaluation of candidate pathophysiological mechanisms and emerging biomarkers.<sup>14</sup>

The aim of this study was therefore to introduce and validate a patient symptom-logging system to aid patients and clinicians in the accurate reporting of gastroduodenal symptoms, including the use of pictograms. Once conceptualized, the standardized patient

**Key points**

- Functional gastroduodenal conditions have overlapping symptomatology which make them difficult to distinguish.
- We developed a pictogram-based symptom-reporting App to improve patient reporting and understanding of gastroduodenal symptoms.
- App symptom scores and PAGI-SYM scores were positively correlated (convergent validity). App symptom scores were negatively correlated with PAGI-QOL scores (concurrent validity). Patient interviews revealed adequate face and content validity.
- The system and App is a valid and patient-centric approach to capturing gastroduodenal symptom experience, which is beneficial for clinical practice and research.

symptom-logging system was implemented in an iOS App and cloud-based reporting portal, before assessment of the convergent, concurrent, face and content validity in patient cohorts.

**2 | MATERIAL AND METHODS**

Ethics approval was granted by the Auckland Health Research Ethics Committee and the Conjoint Health Research Ethics Board at Calgary. All patients provided written informed consent. The study was reported per the Strengthening the Reporting of Observational studies in Epidemiology (STROBE) statement and Standards for Reporting Qualitative Research (SRQR).<sup>15,16</sup>

**2.1 | Gastroduodenal symptom reporting system**

A gastroduodenal symptom logging system was conceptualized as depicted in Figure 1. Ten symptoms were selected for logging based on those covering the spectrum of functional gastroduodenal disorders<sup>4,17</sup>; that is, epigastric pain, epigastric burning, early satiation, excessive post-prandial fullness, nausea, vomiting, bloating, belching, heartburn, and reflux. Symptoms were divided into those that are continuously experienced vs discrete events (Figure 1A). The design required patients to log symptoms at minimum 15-minute intervals, or more frequently if symptoms changed, including around a test meal. Intervals of 15 min have previously been established

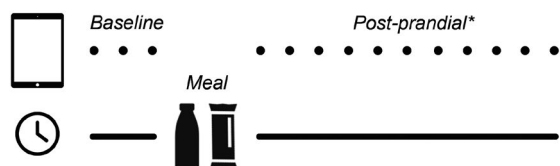
## (A) Inputs

<p><b>'Events'</b></p> <ul style="list-style-type: none"> <li>• Vomit</li> <li>• Reflux</li> <li>• Belch</li> </ul>	<p><b>'Continuous'</b></p> <ul style="list-style-type: none"> <li>• Nausea</li> <li>• Bloating</li> <li>• Heartburn</li> </ul>	<ul style="list-style-type: none"> <li>• Upper Gut Pain</li> <li>• Stomach Burn</li> <li>• Excessive Fullness*</li> </ul>
---	--	---

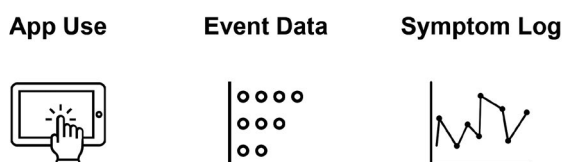
**Single time point:** Early Satiation\*

## (B) App Implementation

Notifications at 15-minute intervals



## (C) Outputs



**FIGURE 1** Pictorial depiction of the gastroduodenal symptom-logging system. \*Refers to symptoms logged only after a test meal

as sufficient to generate accurate gastroduodenal symptom profiling<sup>18,19</sup> (Figure 1B). Discrete symptom events and continuous symptom data were reported graphically (Figure 1C).

## 2.2 | App implementation

After the system was conceptualized, a custom App was implemented in iOS using the Swift 5 programming language, being designed to run on an iPad mini (Apple). The App was developed by Alimetry. The App allowed users to define fasted and/or fed testing durations and to specify a test meal if desired. Symptom reporting was standardized using both pictograms and written descriptors. A pictogram was assigned to each gastroduodenal symptom, being modified from those previously validated in a Belgian FD cohort by Tack et al<sup>9</sup> (Figure 2A and Figure S1). The modifications were undertaken by designers with oversight from clinicians working in the field of GI motility to maintain content validity. The written descriptors were designed to be brief, jargon-free, and employed commonly accepted clinical terminology (Figure 2B). Epigastric pain was simplified to “upper abdominal pain,” and epigastric burning to “stomach burn.”

The severity of each continuous symptom was assessed using a 0–10 Likert scale, with anchors at 0 “none,” indicating no symptom experience, and 10 indicating the “most severe imaginable”

extent of a symptom experience (Figure 2C). This scale was chosen based on recommended guidance from the FDA,<sup>20</sup> and because this scale is sensitive to clinically relevant changes in chronic pain intensity.<sup>21</sup> Events were assessed using a two-step logging interface, allowing users to enter a type of event (Figure 2A), then the timing (Figure 2D). Excessive fullness was only assessed postprandially, while early satiation was only assessed at a single time-point immediately following the meal (Figure 2E). The App displayed notifications every 15 minutes to alert the user to update their symptoms (Figure 2F). Interactions with the App were also continuously tracked to ensure symptom logging completeness and compliance. At the end of the test, the data were automatically transferred from the iPad mini to a Health Insurance Portability and Accountability Act (HIPAA) compliant cloud server (Alimetry; Auckland, New Zealand), for automated generation of a graphical report in a secure online portal (Figure 3).

## 2.3 | Validation study design

Two validation studies were performed to assess the validity of the completed continuous symptom-logging App and its pictograms. Convergent and concurrent validity were evaluated in a multi-center, observational cohort study of patients with chronic gastroduodenal symptoms by comparison with the validated PAGI-SYM and PAGI-QOL instruments. It was hypothesized that the App data would show positive correlations with the PAGI-SYM longer-term measure of gastric symptomatology (convergent validity), whereas for concurrent validity, it was expected that the App data would be negatively correlated with the PAGI-QOL. Semi-structured patient interviews were used to assess face and content validity of the pictograms.

### 2.3.1 | Study 1: Convergent and concurrent validation study

Patients suffering chronic gastroduodenal symptoms were recruited from outpatient services or gastric scintigraphy referral lists. Patients recruited from outpatient services were referred with a diagnosis of gastroparesis, CNVS, or FD, while patients from scintigraphy referrals had chronic gastroduodenal symptoms without further differentiation, thereby ensuring the inclusion of a broad subset of eligible symptomatic patients typical of real-world clinical practice. Participants were excluded if they were aged <18 years, pregnant, or had an identified organic cause for their symptoms including metabolic or endocrine disorders, active GI infections, inflammatory bowel disease, or GI malignancy. All medications known to modify gastrointestinal motility were withheld for 48 hours prior to the study. A correlation coefficient of  $r > 0.3$  was chosen as reasonably indicating validity, and a power calculation showed that a sample size of 79 participants would be needed to detect a difference in the App and the PAGI-SYM and PAGI-QoL outcomes with 80% power, a significance level of 0.05, and an effect size of  $r = 0.31$ .<sup>22</sup>



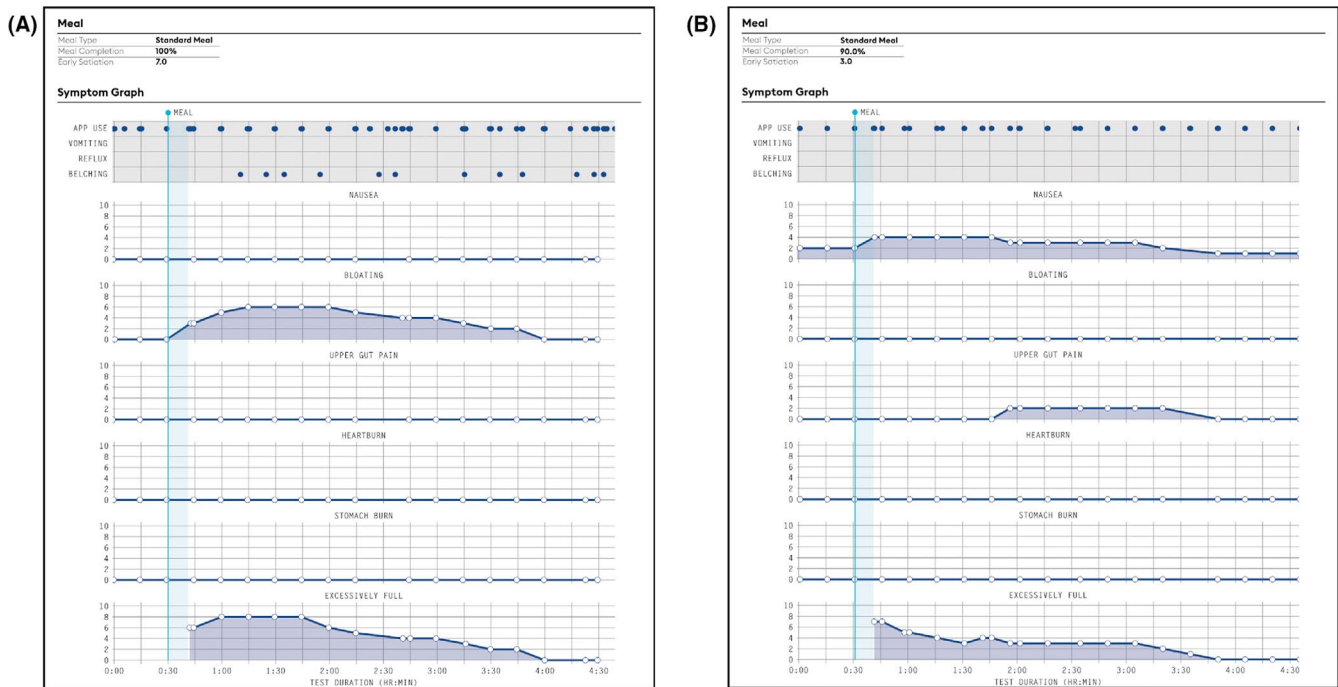
**FIGURE 2** Screenshots of the iOS Symptom-Logging App. (A) Post-meal symptom dashboard display; (B) Symptom explanations display; (C) Upper gut gain symptom update display; (D) Vomiting event logging display; (E) Early satiety symptom update display; and (F) Symptom update reminder display. Content © Alimetry Ltd 2020, provided with permission

The study protocol comprised a 30-min fasted period, followed by consumption of a standardized meal over 10 min, followed by a 4-h post-prandial symptom-logging period. Patients referred from clinics received a nutrient drink (230 ml Ensure; Abbott Nutrition) and an oatmeal energy bar (250 kcal with 5 g fat, 45 g carbohydrate, 10 g protein, 7 g fibre; Clif Bar & Company), whereas patients recruited from scintigraphy lists received a standard egg meal (255 kcal with 72% carbohydrate, 24% protein, 2% fat, 2% fibre) or tofu equivalent if they had an egg allergy. All patients underwent a minimum six-hour pre-test fast.

The PAGI-SYM and PAGI-QOL were completed immediately prior to the start of the 30-minute fasted period. The PAGI-SYM is a 20-item validated questionnaire that asks participants to recall the severity of symptoms experienced over the last 2 weeks using a 6-point Likert scale from “none” to “very severe”.<sup>11</sup> The PAGI-SYM consists of 6 subscales: (i) heartburn/regurgitation; (ii) post-prandial fullness/early satiety; (iii) bloating; (iv) nausea/vomiting; (v) lower abdominal pain; and (vi) upper abdominal

pain. Higher scores reflect higher symptom severity and burden.<sup>11</sup> The PAGI-QoL is a disease-specific health-related quality of life patient-reported outcome measure.<sup>23</sup> It consists of 30 items and 5 subscales: (i) daily activities; (ii) clothing; (iii) diet and food habits; (iv) relationships; and (v) psychological wellbeing and distress.<sup>23</sup> It also asks participants to recall the previous 2 weeks’ experiences and uses a 6-point Likert scale. Throughout the study period, participants used the App to log their symptoms as described above. At the end of the study, participants also completed a single 5-point Likert scale to assess the ease-of-use of the App (0 = Very easy; 1 = somewhat easy; 2 = neutral; 3 = somewhat difficult; and 4 = very difficult).

Statistical analyses were performed in GraphPad Prism v.9.1.2 (GraphPad). The PAGI-SYM summary score was calculated by taking the mean of all subscale scores.<sup>11</sup> The PAGI-QOL summary score was calculated as the mean of all subscales, when all item responses were reverse coded,<sup>23</sup> such that a higher PAGI-QOL score represents better disease-specific quality of life. Participant engagement with the



**FIGURE 3** Examples of gastric test symptom reports following meal consumption in (A) A participant with functional dyspepsia and (B) A participant with chronic nausea and vomiting syndrome. Content © Alimetry Ltd 2020, provided with permission

App was assessed as the average time elapsed between App interactions, with non-compliance defined as median time exceeding 30 min between symptom logs. Before coding the App, a test was conducted on 5 patients with the symptom-logging screens printed on paper to confirm that the system was working appropriately. Data from these patients were therefore excluded from the App usability and compliance testing data. Symptom-logging data was taken from the 4-h post-prandial period to reflect when participants most likely experience symptoms in their everyday lives.<sup>24</sup> The mean score and the area under the curve (AUC) were calculated for each symptom that was continuously logged during the 4-h post-prandial period, or as a single time-point for early satiation. A total symptom burden (referred to as the “Gastric Symptom Burden Score”) was also calculated, as both the sum of each participant's mean symptom scores with early satiation, and as the sum of individual symptom AUCs excluding early satiation. Spearman's correlations were assessed between the Gastric Symptom Burden score obtained using each of these two methods. Discrete “events” captured by the App (vomiting, reflux, and belching) were not included in the individual symptom analysis nor the Gastric Symptom Burden score.

Spearman's correlations were calculated to assess convergent validity between the post-prandial App-derived symptom severity scores (early satiation, bloating, upper gut pain, heartburn, excessive fullness, and nausea) and the outcome of the corresponding PAGI-SYM item. Spearman's correlation was also used to assess the association between the Gastric Symptom Burden score and the PAGI-SYM summary score (convergent validity) and gastric-specific quality of life as measured by the PAGI-QOL (concurrent validity).

$p < 0.05$  was considered statistically significant. Weak, moderate, and strong correlations were defined as  $r$  values greater than 0.1, 0.3, and 0.5, respectively.

### 2.3.2 | Study 2: Face and content validation study

Adults aged  $\geq 18$  years with a diagnosis of gastroparesis, CNVS, or FD (as per Rome IV), who resided in New Zealand and were able to provide informed consent, were eligible for the qualitative interview study. Potential participants were recruited via social media advertising, patient peer support groups, and clinical referrals. Exclusion criteria included the inability to speak or read English, and vulnerable participants (e.g., prisoners, individuals with a known cognitive impairment). Recruitment and data collection occurred between June 2020 and July 2021. We aimed to recruit between 5 and 15 participants in line with previous pictogram validation studies.<sup>9,25</sup>

Semi-structured interviews were conducted using a web-based conferencing platform by two researchers trained and experienced in qualitative research methods. Participants were invited to have a support person present and interviews were scheduled according to participant request to ensure optimal conditions including privacy. During the interview, one pictogram was presented at a time and participants were asked to describe the symptom that was represented by the pictogram. After all pictograms had been presented, the exercise was repeated but participants were asked to choose the symptom label that they thought best represented the symptom depicted in the pictogram from a prescribed list. Finally, there was a

general discussion regarding any pictograms that were deemed to be poor representations, any symptoms that had not been well represented, and any recommendations for improvement. The interviews were recorded and transcribed by one of the interviewers with participant consent. Missing responses were excluded from analysis.

An adapted form of iterative thematic analysis, with a semi-quantitative approach, was utilized to evaluate participant responses. Agreement between the participants' interpretation of the symptom pictogram and the "intended symptom" was calculated and is reported as a percentage. Participant commentary was coded for each pictogram and analyzed for common themes. Two coders, including one who did not participate in the interviews, coded all participant responses based on *a priori* coding categories. Coding agreement was assessed after transcription of the first two interviews and after coding of all interviews.

### 3 | RESULTS

#### 3.1 | Study 1: Convergent and concurrent validity

A total of 79 patients were included with median age 41 years (IQR = 27–52), most being female (82%), and a majority self-identifying as Caucasian (71%). All tests except one were performed in the morning. Two subjects who completed the face and content validity study below were subsequently also recruited into the convergent and concurrent validity study. Table 1 provides a detailed description of participant characteristics. Patients had a representative mix of FD (12.7%), CNVS or gastroparesis (53.2%), and chronic gastroduodenal symptoms not further delineated (scintigraphy referral group; 34.2%). Across the whole cohort, 16.5% had a concurrent diagnosis of irritable bowel syndrome, and 26.6% had a concurrent psychological comorbidity (anxiety, depression, and/or post-traumatic stress syndrome).

On average, participants interacted with the App at a median of 8.5-minute intervals (IQR = 6.5–10.2), indicating high engagement. The average time between symptom logs did not exceed 13.6 min in any patient, demonstrating a 100% compliance rate with symptom logging. Of all participants who used the app, 90.1% reported that it was "very easy to use" and 9.9% reported it was "somewhat easy". Three participants did not respond to the ease-of-use question.

Example symptom report outputs are provided in Figure 3A,B for patients with FD and CNVS, respectively. Across all subjects, strong correlations were found between the mean and AUC metrics for the App post-prandial symptom scores, and for the calculated Gastric Symptom Burden Score ( $r_s = 0.95$ – $0.99$ ). The validity comparisons therefore only employed the mean scores, with the AUC metrics not being further employed.

Moderate to strong associations were found for all assessed symptoms ( $p < 0.001$ ) between the App and PAGI-SYM individual symptom measures (Figure 4). Early satiation ( $r_s = 0.55$ ) and nausea ( $r_s = 0.67$ ) correlated strongly, while upper gut pain ( $r_s = 0.40$ ), heartburn ( $r_s = 0.47$ ), bloating ( $r_s = 0.42$ ), and excessive fullness ( $r_s = 0.40$ )

demonstrated moderate associations with their PAGI-SYM equivalent (Figure 4). There was a strong association between the total Gastric Symptom Burden Score and the mean PAGI-SYM summary score ( $r_s = 0.56$ ,  $p < 0.001$ ) (Figure 5A). These results confirm convergent validity between the App and the PAGI-SYM questionnaire.

The Gastric Symptom Burden Score had a significant, moderate, and negative association with the PAGI-QOL summary score ( $r_s = -0.34$ ,  $p = 0.002$ ) (Figure 5B). This finding confirms concurrent validity, as participants who reported a higher burden of post-prandial gastric symptoms in the App were also likely to report reduced quality of life.

#### 3.2 | Study 2: Face and content validity

Eight patients participated in the pictogram validation interviews. However, one interview was excluded due to loss of data, and a second was excluded as the participant had completed the convergent and concurrent validity study prior to their interview and were therefore not pictogram naive. The analyzed interview cohort consisted of Caucasian females of median age 38 years (IQR = 25.5–46.0), all with CNVS or gastroparesis. The interview length ranged from 9 to 36 min.

The highest performing pictogram was that depicting vomiting, which had 100% agreement, meaning all participants suggested that the pictogram described vomiting as it was intended to do. Pictograms for belching (83% agreement), upper abdominal pain (83%), heartburn (67%), bloating (67%), nausea (67%), and reflux (67%) showed high to moderate degrees of agreement. Early satiation (0%) and excessive fullness (16%) had poor participant agreement.

When participants were provided with the list of symptoms to match with the pictogram, agreement improved. Belching, heartburn, nausea, reflux, upper abdominal pain, and vomiting showed 100% agreement. Feeling excessively full (80% agreement), bloating (80%), and early satiation (50%) had strong to moderate agreement.

Assessment of patient verbal assessments of the pictograms contributed little further data beyond the above analyses, generally matching the summarized findings. The vomiting pictogram received the most positive feedback, whereas early satiation received the most negative feedback with four out of six participants stating that the pictogram did not match well with the symptom label. Suggestions regarding other pictograms (feeling excessively full, heartburn, nausea, and reflux) were heterogenous (example responses provided in Table S1).

### 4 | DISCUSSION

This study reports the design, implementation, and validation of a standardized system, App, and cloud-based reporting portal for continuous patient symptom logging in functional gastroduodenal disorders. The App, coded in iOS, differentiates continuous

TABLE 1 Participant clinical and demographic characteristics

Characteristic	FD (n = 10)	CNVS or Gastroparesis (n = 42)	Scintigraphy (n = 27)	Total sample (N = 79)
Age (years; median) (IQR)	51 (32–64)	34 (26–44)	45 (31–56)	41 (27–52)
Female, n (%)	9 (90)	37 (88)	19 (70)	65 (82)
Ethnicity, n (%)				
Caucasian	9 (90)	31 (74)	16 (59)	56 (71)
Māori	0 (0)	3 (7)	4 (15)	7 (9)
Chinese	1 (10)	1 (2)	1 (4)	3 (4)
Pasifika	0 (0)	1 (2)	1 (4)	2 (3)
Indian	0 (0)	1 (2)	1 (4)	2 (3)
Other	0 (0)	5 (12)	4 (15)	9 (11)
BMI (kgm <sup>2</sup> ) mean (SD)	22.6 (4.8)	24.1 (4.4)	24.0 (4.5)	23.9 (4.5)
Comorbidities, n (%)				
Diabetes	1(10)	5 (12)	1 (4)	7 (9)
Hypo/Hyper-thyroidism	1(10)	0 (0)	1 (4)	2 (3)
Cardio-Respiratory	1(10)	8 (19)	6 (22)	15 (19)
IBS	1(10)	7 (17)	5 (19)	13 (16)
Anxiety/Depression/PTSD	2 (20)	16 (38)	3 (11)	21 (27)
PAGI-SYM Score Mdn (IQR)	1.23 (1.03–1.73)	2.45 (1.83–3.30)	2.60 (1.40–3.45)	2.40 (1.50–3.30)
PAGI-QOL Score Mdn (IQR)	3.34 (2.74–4.39)	2.68 (1.82– 3.44)	3.50 (2.40–3.80)	2.86 (2.03–3.71)
GCSI Score Mdn (IQR)	1.71 (1.31–2.47)	3.17 (2.52–3.67)	2.83 (2.08–3.64)	2.83 (2.06–3.64)

Abbreviations: BMI, Body Mass Index; GCSI, Gastroparesis Cardinal Symptom Index; IBS, Irritable Bowel Syndrome; PTSD, post-traumatic stress disorder.

symptoms from discrete symptom events and incorporates standardized symptom identifiers and scales for severity gradings. Robust correlations were shown between the App and PAGI-SYM measures of individual and overall gastric symptoms indicating convergent validity. The association between PAGI-QOL scores and App-reported symptoms reflected an expected relationship between gastric-specific quality of life and overall gastric symptom burden, demonstrating concurrent validity. Most patients reported the App as “very easy to use,” validating usability. Qualitative interviews demonstrated adequate face and content validity of the pictogram-based approach to symptom reporting, although finding some areas for improvement. Overall, these findings show that the App-based 4-hour “symptom snapshot” following a standard meal is a valid, patient-centric, and representative method for evaluating gastroduodenal symptoms.

Previous literature has identified pictograms as an efficient way of communicating subjective symptoms, being superior to text alone,<sup>9,25,26</sup> particularly in conditions with frequently overlapping symptoms where verbal descriptors do not allow all patients to contextualize the individual and multi-dimensional nature of their symptoms.<sup>9</sup> Tack et al<sup>9</sup> previously introduced a series of pictograms specific for functional GI disorders and demonstrated improved patient symptom reporting accuracy and symptom understanding in a cohort of FD patients. In the present study, these pictograms were modified and incorporated into our App to support standardized real-time symptom capture around a meal. Our qualitative data showed that the depiction of vomiting performed ideally, and those

for belching, heartburn, upper abdominal pain, and bloating also performed strongly. However, further improvements are desirable to better depict early satiation and excessive fullness in subsequent iterations of the App.

A recent study by Kuwelker et al<sup>19</sup> similarly reported that symptoms captured throughout a 4-h gastric scintigraphy study correlated well with validated measures of gastric symptoms in patients with diabetes mellitus. As such, the findings from the current study corroborate the validity of the 4-h gastric symptom capture window and extend these findings to a greater range of patients with gastroduodenal symptoms.<sup>19</sup> Together, these two studies provide a strong foundation for the external validity of utilizing a 4-h gastric symptom snapshot to assess gastric symptomatology in clinical practice and research.<sup>19</sup>

The PAGI-SYM questionnaire is one of the few generic functional gastroduodenal symptom severity metrics which has good test-retest consistency, is well-validated, and is easily administered.<sup>11</sup> The GCSI Daily Dairy, employing a key subset of the PAGI-SYM items, was subsequently introduced as a validated method of daily symptom reporting.<sup>12</sup> A small number of other tools exist, but their uptake is disease-specific or not widespread.<sup>12,27,28</sup> There is a paucity of standardized and validated techniques to enable the assessment of gastric symptoms continuously over a defined test period. This App will therefore fill a significant gap, offering broad utility. In clinical settings, it offers standardized accurate symptom reporting to aid patient-clinician communication, being particularly useful when symptom correlations are required for diagnostic tests

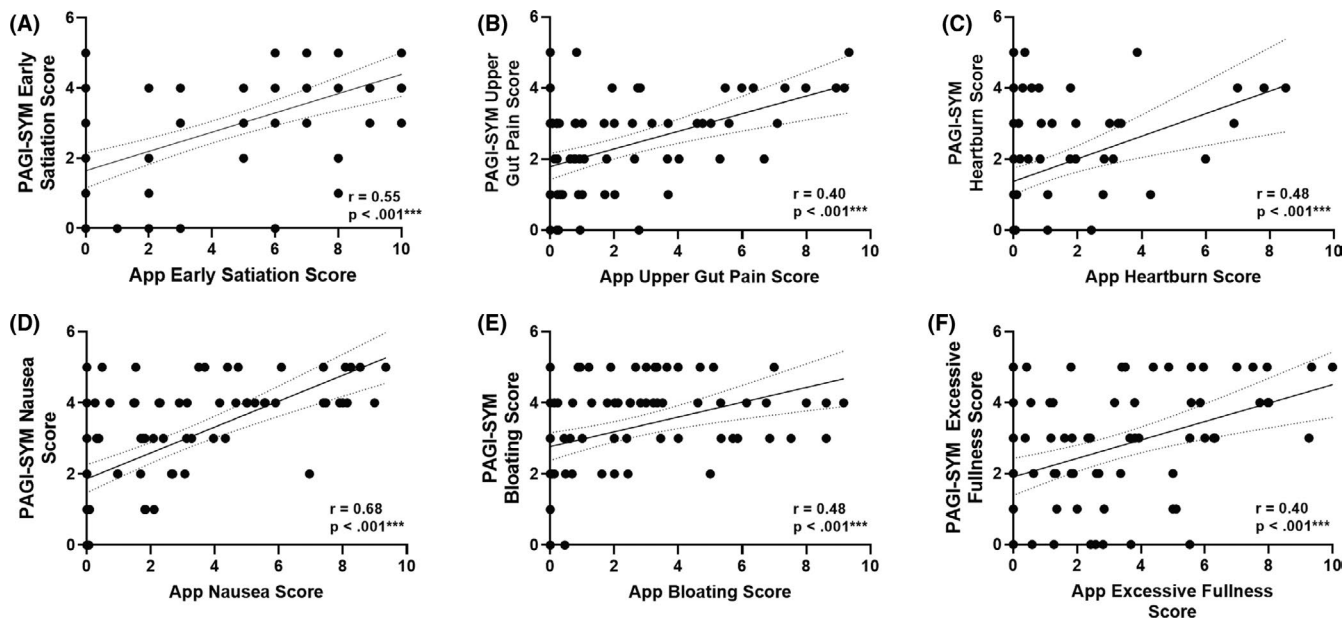


FIGURE 4 Scatterplots showing correlations between App symptom scores and related PAGA-SYM item scores ( $n = 79$ ). Dotted lines represent the corresponding 95% confidence intervals

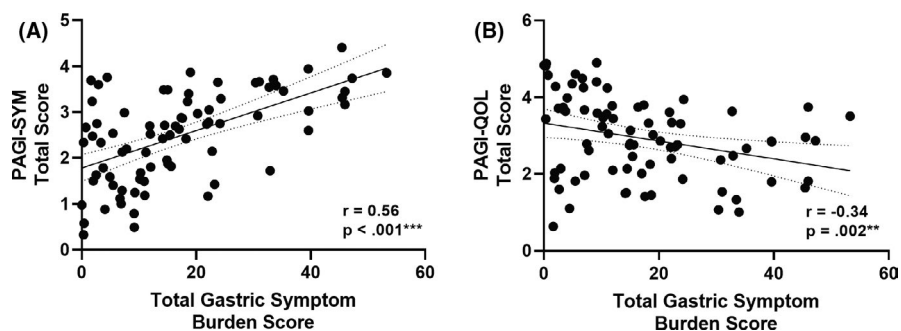


FIGURE 5 Scatterplots showing correlations between the total Gastric Symptom Burden Score reported via the App vs. (A) PAGA-SYM overall score and (B) PAGA-QOL overall score ( $n = 79$ ). Dotted lines represent the corresponding 95% confidence intervals

involving provocations, such as nutrient drink tests of accommodation, gastric emptying tests, and body surface gastric mapping.<sup>19,29,30</sup> The symptoms captured in the App were derived from accepted standards in gastroenterology, and in particular the cardinal symptoms of gastroparesis and Rome IV criteria for, chronic nausea and vomiting syndromes and functional dyspepsia (including epigastric pain syndrome and post-prandial distress subtypes), as well as their commonly-associated features such as belching and reflux.<sup>2,4,6,9</sup> The App is also anticipated to be useful in defining and comparing symptom experiences before and after interventions, and for pathophysiological studies employing the “Plausibility Criteria” for mechanisms of functional GI disorders.<sup>15,20</sup>

Some limitations to this study are acknowledged. While a relatively large multi-center, international, prospective cohort of symptomatic patients were included, specific patient subgroups entering the study via scintigraphy referral lists were not completely defined by Rome Criteria or other diagnoses. This allowed demonstration of validity over a diverse patient cohort but precluded the assessment of criterion validity. Symptom correlations with PAGA-SYM were only possible for the continuous symptoms, and not the discrete

symptom events. It is also acknowledged that while the 4-h symptoms snapshot may be representative of an individual's PAGA-SYM responses, this tool is not designed to capture the comprehensive and nuanced fluctuations of symptoms over longer time periods. Two test meals were included, enabling validity to be demonstrated across a range of meal types; however, this may have diversified symptom profiling during the post-prandial window. Additionally, within our cohort, all but one patient was studied in the morning, whereas recent data suggests that gastric symptom severity may be at its worst in the evening in gastroparesis patients.<sup>31</sup> It is notable that despite these factors, symptom reporting via the App still demonstrated robust validity in capturing a reasonably typical symptom burden experience.

A limitation to the qualitative study was that the pictogram validation cohort consisted only of Caucasian women with CNVS or gastroparesis. Such patients are typical of CNVS and gastroparesis demographics in our practice; however, culture, gender, and disorder experiences could all possibly influence symptom interpretations.<sup>32-35</sup> Self-reporting, recall, and social desirability biases are also possible due to the inherent nature of interview data.<sup>36,37</sup>



Nevertheless, the interviews revealed consistent and helpful insights including areas for improvement. The gastroduodenal symptom-logging App will be iteratively improved over time, presenting future opportunities for re-evaluation in an increasingly diverse interview group. Future iterations may optimize usability for at-home use or long-term monitoring, for example, by changing the notification scheme, being outside the current scope of the App.

In summary, this study introduces a continuous patient symptom-logging system, App, and reporting tool for gastroduodenal disorders, and demonstrates that a 4-hour post-prandial symptom snapshot is a valid and representative approach to measuring symptom experiences and their burden. The system and App are anticipated to be a useful addition to clinical practice and research.

## ACKNOWLEDGEMENTS

The authors would like to thank the patients who took part in this study and our clinical recruitment staff Gen Johnston, India Wallace, Lynn Wilsack and Renata Rehak. We thank Jim Hannon-Tan and Tom Russell for their invaluable input to the App and pictogram design. The validation studies were supported by the Health Research Council of New Zealand and the John Mitchell Crouch Fellowship from the Royal Australasian College of Surgeons. Open access publishing facilitated by The University of Auckland, as part of the Wiley - The University of Auckland agreement via the Council of Australian University Librarians.

## CONFLICT OF INTEREST

GOG and AAG hold grants and intellectual property in the field of gastrointestinal electrophysiology and are Directors in Alimetry. GOG is a Director in The Insides Company. PD is a Director in FlexiMap. SC, GS, CK, JSTW, SJW PD and CNA are members of Alimetry. The remaining authors have no conflicts of interest to declare.

## AUTHOR CONTRIBUTIONS

GS, NK, GOG, AAG, EB, and CK were involved in study conception and design. GS, NK, CK and DC, were involved in data collection. GS, NK, SC, JSTW, CK, DC, CV, PD, SJW, JT, CNA, EB, AAG and GOG were involved in the data analysis and interpretation. GS, CV, GOG, CK were involved in drafting the manuscript. GS, NK, SC, JSTW, CK, DC, CV, PD, SJW, JT, CNA, EB, AAG, GOG were involved in critical revisions/final approval of the manuscript.

## ORCID

Gabrielle Sebaratnam  <https://orcid.org/0000-0002-9946-9815>

Nikita Karulkar  <https://orcid.org/0000-0001-9177-0382>

Stefan Calder  <https://orcid.org/0000-0002-5855-1732>

Celia Keane  <https://orcid.org/0000-0002-9611-2335>

Daniel A. Carson  <https://orcid.org/0000-0002-5089-2807>

Chris Varghese  <https://orcid.org/0000-0001-7369-8639>

Peng Du  <https://orcid.org/0000-0002-6913-7545>

Stephen J. Waite  <https://orcid.org/0000-0002-2856-0566>

Jan Tack  <https://orcid.org/0000-0002-3206-6704>

Elizabeth Broadbent  <http://orcid.org/0000-0003-3626-9100>

Armen A. Gharibans  <http://orcid.org/0000-0002-9139-4740>

Greg O'Grady  <http://orcid.org/0000-0002-5998-1080>

## REFERENCES

1. Sperber AD, Bangdiwala SI, Drossman DA, et al. Worldwide prevalence and burden of functional gastrointestinal disorders, results of Rome foundation global study. *Gastroenterology*. 2021;160(1):99-114.
2. Lacy BE, Parkman HP, Camilleri M. Chronic nausea and vomiting: evaluation and treatment. *Am J Gastroenterol*. 2018;113(5):647-659.
3. Aro P, Talley NJ, Agréus L, et al. Functional dyspepsia impairs quality of life in the adult population. *Aliment Pharmacol Ther*. 2011;33(11):1215-1224.
4. Stanghellini V, Chan FKL, Hasler WL, et al. Gastroduodenal disorders. *Gastroenterology*. 2016;150(6):1380-1392.
5. Harer KN, Pasricha PJ. Chronic unexplained nausea and vomiting or gastric neuromuscular dysfunction (GND)? An update on nomenclature, pathophysiology and treatment, and relationship to gastroparesis. *Curr Treat Options Gastroenterol*. 2016;14(4):410-419.
6. Pasricha PJ, Colvin R, Yates K, et al. Characteristics of patients with chronic unexplained nausea and vomiting and normal gastric emptying. *Clin Gastroenterol Hepatol*. 2011;9(7):567-576.
7. Geeraerts A, Van Houtte B, Clevers E, et al. Gastroesophageal reflux disease-functional dyspepsia overlap: do birds of a feather flock together? *Am J Gastroenterol*. 2020;115(8):1167-1182.
8. von Wulffen M, Talley NJ, Hammer J, et al. Overlap of irritable bowel syndrome and functional dyspepsia in the clinical setting: prevalence and risk factors. *Dig Dis Sci*. 2019;64(2):480-486.
9. Tack J, Carbone F, Holvoet L, Vanheel H, Vanuytsel T, Vandenbergh A. The use of pictograms improves symptom evaluation by patients with functional dyspepsia. *Aliment Pharmacol Ther*. 2014;40(5):523-530.
10. Zhao W, Jin H, Zhang L, et al. Pictograms to provide a better understanding of gastroesophageal reflux symptoms in chinese subjects. *Gastroenterol Res Pract*. 2017;2017:1214584.
11. Rentz AM, Kahrilas P, Stanghellini V, et al. 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(10):1737-1749.
12. Revicki DA, Rentz AM, Dubois D, et al. Gastroparesis Cardinal Symptom Index (GCSI): development and validation of a patient reported assessment of severity of gastroparesis symptoms. *Qual Life Res*. 2004;13(4):833-844.
13. Revicki DA, Camilleri M, Kuo B, et al. Development and content validity of a gastroparesis cardinal symptom index daily diary. *Aliment Pharmacol Ther*. 2009;30(6):670-680.
14. Tack J, Corsetti M, Camilleri M, et al. Plausibility criteria for putative pathophysiological mechanisms in functional gastrointestinal disorders: a consensus of experts. *Gut*. 2018;67(8):1425-1433.
15. von Elm E, Altman DG, Egger M, et al. Strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *BMJ*. 2007;335(7624):806-808.
16. O'Brien BC, Harris IB, Beckman TJ, Reed DA, Cook DA. Standards for reporting qualitative research: a synthesis of recommendations. *Acad Med*. 2014;89(9):1245-1251.
17. Camilleri M, Chedid V, Ford AC, et al. Gastroparesis. *Nat Rev Dis Primers*. 2018;4(1):41.
18. Bisschops R, Karamanolis G, Arts J, et al. Relationship between symptoms and ingestion of a meal in functional dyspepsia. *Gut*. 2008;57(11):1495-1503.
19. Kuwelker S, Prichard DO, Bailey K, Bharucha AE. Relationship between symptoms during a gastric emptying study, daily

- symptoms and quality of life in patients with diabetes mellitus. *Neurogastroenterol Motil.* 2021;33(12):e14154.
20. U.S. Food and Drug Administration. Gastroparesis: Clinical Evaluation of Drugs for Treatment Guidance for Industry [Internet]. 2015 [cited 2021 Aug 16]. Available from: <https://www.fda.gov/files/drugs/published/Gastroparesis—Clinical-Evaluation-of-Drugs-for-Treatment.pdf>
  21. Farrar JT, Young JP Jr, LaMoreaux L, Werth JL, Poole MR. Clinical importance of changes in chronic pain intensity measured on an 11-point numerical pain rating scale. *Pain.* 2001;94(2):149-158.
  22. Faul F, Erdfelder E, Buchner A, Lang A-G. Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behav Res Methods.* 2009;41(4):1149-1160.
  23. De la loge C, Trudeau E, Marquis P, et al. Cross-cultural development and validation of a patient self-administered questionnaire to assess quality of life in upper gastrointestinal disorders: the PAGI-QOL. *Qual Life Res.* 2004;13(10):1751-1762.
  24. Castillo E, Camilleri M, Locke III G, et al. A community-based, controlled study of the epidemiology and pathophysiology of dyspepsia. *Clin Gastroenterol Hepatol.* 2004;2(11):985-996.
  25. Beng TS, Ann YH, Guan NC, et al. The suffering pictogram: measuring suffering in palliative care. *J Palliat Med.* 2017;20(8):869-874.
  26. Dowse R. Pharmacists, are words enough? The case for pictograms as a valuable communication tool. *Res Social Adm Pharm.* 2021;17(8):1518-1522.
  27. Chan Y, Cheong PK, Fang FF, et al. A symptom severity questionnaire for patients suffering from functional gastrointestinal disorder: FGI-Checklist. *J Gastroenterol Hepatol.* 2020;35(7):1130-1135.
  28. Adam B, Liebrechts 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(4):357-363.
  29. Boeckstaens GE, Hirsch DP, van den Elzen BD, Heisterkamp SH, Tytgat GN. Impaired drinking capacity in patients with functional dyspepsia: relationship with proximal stomach function. *Gastroenterology.* 2001;121(5):1054-1063.
  30. Carson DA, O'Grady G, Du P, Gharibans AA, Andrews CN. Body surface mapping of the stomach: new directions for clinically evaluating gastric electrical activity. *Neurogastroenterol Motil.* 2021;33(3):e14048.
  31. Shahsavari D, Yu D, Jehangir A, Lu X, Zoll B, Parkman HP. Symptom variability throughout the day in patients with gastroparesis. *Neurogastroenterol Motil.* 2020;32(2):e13740.
  32. Chang L, Toner BB, Fukudo S, et al. Gender, Age, Society, Culture, and the Patient's Perspective in the Functional Gastrointestinal Disorders. *Gastroenterology.* 2006;130:1435-1446.
  33. López-Colombo A, Morgan D, Bravo-González D, Montiel-Jarquín A, Méndez-Martínez S, Schmulson M. The epidemiology of functional gastrointestinal disorders in Mexico: a population-based study. *Gastroenterol Res Pract.* 2012;2012:606174.
  34. Fang X, Francisconi CF, Fukudo S, et al. Multicultural aspects in functional gastrointestinal disorders (FGIDs). *Gastroenterology.* 2016;150(6):1344-1354.
  35. Lee H-J, Lee S-Y, Kim JH, et al. Depressive mood and quality of life in functional gastrointestinal disorders: differences between functional dyspepsia, irritable bowel syndrome and overlap syndrome. *Gen Hosp Psychiatry.* 2010;32(5):499-502.
  36. Bauhoff S. Systematic self-report bias in health data: impact on estimating cross-sectional and treatment effects. *Health Serv Outcomes Res Methodol.* 2011;11(1):44-53.
  37. Shrout PE. Commentary: mediation analysis, causal process, and cross-sectional data. *Multivariate Behav Res.* 2011;46(5):852-860.

### SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Sebaratnam G, Karulkar N, Calder S, et al. Standardized system and App for continuous patient symptom logging in gastroduodenal disorders: Design, implementation, and validation. *Neurogastroenterology & Motility.* 2022;00:e14331. doi:[10.1111/nmo.14331](https://doi.org/10.1111/nmo.14331)