Psychometric development of the Gastrointestinal Symptom Rating Questionnaire (GSRQ) demonstrated good validity

Hayley A Hutchings[1]*, Wai-Yee Cheung[1], Ian T Russell[1], Dharmaraj Durai[2], Laith Alrubaiy[1], John G Williams[1]

[1] HA Hutchings, Associate Professor in Health Services Research

W-Y Cheung, Senior Lecturer in Health Services Research

IT Russell, Professor of Clinical Trials

L Alrubaiy, PhD student

JG Williams, Consultant Gastroenterologist / Professor of Health Services Research

Patient and Population Health and Informatics Research

College of Medicine

Swansea University

Singleton Park

Swansea

SA2 8PP

United Kingdom

[2]D Durai, Consultant Gastroenterologist

University Hospital of Wales

Cardiff and Vale University Health Board

Heath Park

Cardiff

CF14 4XW

*Corresponding author

h.a.hutchings@swansea.ac.uk

Telephone: +44 (0) 1792 513412

Abstract

Objective

To develop and validate a gastrointestinal (GI) symptom rating questionnaire for patients with luminal gastrointestinal symptoms including where no diagnosis has been made.

Study Design and Setting

We developed and validated the Gastrointestinal symptom rating questionnaire (GSRQ) in three stages: 1) item generation to identify the relevant items for scale inclusion; 2) development and piloting on patients with a known gastrointestinal disorder; and 3) testing in a sample of trial patients. We examined the underlying dimensions of the scale, internal consistency, validity, reproducibility and responsiveness.

Results

We identified four interpretable factors on the GSRQ. The GSRQ had good internal consistency, (corrected item-subscale correlations between 0.4 and 0.8), and Cronbach's alpha greater than 0.7 for each sub-scale. Construct validity was demonstrated by modest but significant correlations with the SF-36 and the EQ5D index value. We demonstrated good reproducibility with intra-class correlations for test retest scores between 0.71 and 0.77, and significant responsiveness ratios for all sub-scales in patients who had improved, and in two of the sub-scales in patients who had deteriorated.

Conclusion

The GSRQ could be a useful tool to monitor quality of life in various luminal gastrointestinal conditions and where a formal diagnosis has not been made.

Key words

quality of life, gastrointestinal symptoms, validation, development, psychometric analysis

Running head

The Gastrointestinal Symptom Rating Questionnaire

What is new?

Key findings

- The GSRQ was successfully validated on patients with no confirmed diagnosis as well as during the course of treatment following diagnosis

What this study adds to what is known

- Valid health related quality of life (HRQL) instruments are needed to assess and monitor patients attending clinics with gastrointestinal (GI) symptoms. Although numerous questionnaires exist to measure HRQL in patients with GI symptoms, there are no validated instruments available for use at first referral when a diagnosis has not been made. The GSRQ has the potential to help monitor HRQL in patients before formal diagnosis and during the longitudinal course of their disease

What is the implication?

- The GSRQ has the potential to help monitor HRQL in patients before formal diagnosis and during the longitudinal course of their disease

What should change now?

- The GSRQ should be routinely used in clinical practice to measure health related quality of life in a variety of luminal disorders

Introduction

Monitoring patient health related quality of life (HRQL) has become a key part of research and health care in recent years. Gastrointestinal (GI) symptoms are common in the adult and elderly population in North America, Europe and the UK[1,2,3]. In Europe the prevalence of upper GI symptoms ranges from 25 to 35% and for lower GI symptoms from 3 to 22%[4,5]. It is estimated that up to 40% of adults in the UK suffer from GI symptoms in any one year[6,7,8]. In addition, around 50% of new referrals to secondary care gastroenterology clinics are patients who present with GI symptoms but no identifiable structural or biochemical abnormalities[9,10]. These GI symptoms adversely affect patients' well-being and their ability to enjoy day-to-day activities[11]. Valid instruments are therefore needed to assess and monitor the progress of patients attending gastroenterology clinics with gastrointestinal symptoms.

There has been some success in using generic instruments such as such as the Short Form 36 (SF36)[12], Psychological General Wellbeing Scale (PGWB)[13] and Sickness Impact Profile (SIP)[14], to assess the health status of GI patients. However, there are concerns that these instruments might miss small but clinically important changes[15].

Since the 1990s, there has been an exponential growth in the number of quality of life measures developed for patients with GI disorders [15,16]. Disease-specific instruments have been developed for inflammatory bowel disease (IBD), irritable bowel syndrome (IBS), dyspepsia, gastroesophageal reflux disease (GORD), liver disease and GI

malignancy[17,18]. There are however, many disorders for which no valid instruments exist. Furthermore, it is not always appropriate to use disease-specific instruments for newly-referred patients who have not yet had a confirmed diagnosis. The EORTC Quality of Life Group (http://groups.eortc.be/qol/) have recently developed quality of life questionnaires for patients with cancer, with modular 'add-ons' for GI disorders, such as the EORTC QLC-GINET21[18]. These may have been a useful starting point for our instrument development, however they were not available when we undertook the study.

Classification and management of GI disorders is often symptom based, and it is common for people with different GI disorders, to have similar symptoms[19]. Disease specific instruments for different GI disorders, even those supposed to be related to different anatomic regions, often ask patients' views on similar symptom groups[15,16]. However, it is doubtful whether disease-specific instruments that assess HRQL in one single GI disorder are appropriate for assessing the health status of all GI patients. GI luminal disorders cover disorders along the entire GI tract from the mouth to the bowel, and in the course of diagnosis and monitoring would require a much broader assessment of symptoms than the specific questions that are asked in disease-specific questionnaires. To capture small but clinically important changes in GI symptoms, an optimum approach for patients with luminal disorders would be to use a system-specific instrument, i.e. one developed for all GI disorders. There are very few of these systemic instruments available. The most well-known, the Gastrointestinal Quality of Life Index

(GIQLI)[20] is validated for use with patients with a confirmed diagnosis. In the course of undertaking evaluations of interventions that were designed to make a diagnosis the research team became aware that there was no validated instrument available that could be used at first referral and be applied to patients where no confirmed diagnosis had been made; and that could subsequently be used longitudinally to follow the course of the disease on the basis of their symptoms. The aim of this study therefore was to develop and validate a symptom-rating questionnaire suitable to measure the health status of patients with luminal GI symptoms referred to secondary care and in particular where a diagnosis has not been established. This publication focuses on the in-depth validation of this new measure- the Gastrointestinal Symptom Rating Questionnaire (GSRQ).

Materials and Methods

We adapted Streiner and Norman's approach[21] to develop the instrument in the following four stages:

Stage 1: Item generation for the Gastrointestinal Symptom Rating Scale (GSRQ)

Stage 2: Pilot study at a local hospital for initial validation

Stage 3: Main study for concurrent validation in the context of a national multi-institution nurse endoscopy trial (MINuET)[22,23].

Stage 1: Item generation

We carried out a detailed review of the literature using the search terminology 'quality of life', 'questionnaire', 'validation' and 'gastroenterology', in order to determine the

most relevant items for a gastrointestinal symptom rating scale. We identified a number of existing questionnaires which contained items that were potentially suitable for inclusion in the development of a new scale (UK Inflammatory Bowel Disease Questionnaire (UK IBDQ)[24], the Aberdeen Dyspepsia Scale (ADS)[25], the Gastro-oesophageal Reflux Disease Health Related Quality of Life Scale (GERD-HRQLS)[26],the Irritable Bowel Syndrome Quality of Life Questionnaire (IBS QOL)[27], the Gastrointestinal Symptom Rating Questionnaire (GSRQ)[28] and the Gastrointestinal Quality of Life Index (GIQLI)[20]).

A panel of clinicians, patients and public with expertise in gastroenterology, psychology, outcome measurement and methodology reviewed these items and developed the initial version of the GSRQ containing 30 items (see Appendix 1). The panel were asked whether they considered that all appropriate symptoms had been included in the questionnaire, that the questions were appropriately worded and that they were suitable for patients with and without a diagnosis of a GI luminal disorder.

In addition, we separately piloted the developed questionnaire with a purposive sample of 10 patients with a confirmed GI condition from a local hospital (Neath PortTalbot, UK). Patients were asked to complete the questionnaire as well as four supplementary questions:

- Did you find any of the questions difficult to understand?
    - Yes/No. If yes, which one(s) and why?

- Was there any question you did not want to answer?
    - Yes/No. If yes, which one(s) and why?
- Were there any specific aspects of your bowel condition that were not covered by these questions?
    - Yes/No. If yes, which one(s) and why?
- Did you find any of these questions not applicable to you?
    - Yes/No. If yes, which one(s) and why?

The questions were organized into hypothesized domains on the basis of similarity of symptoms.

Stage 2: Pilot study initial validation

We carried out an initial validation of the GSRQ on a sample of patients with a known luminal GI disorder (Dyspepsia, GORD, IBD and IBS) at a local hospital (Neath PortTalbot, UK). We invited patients to complete a questionnaire containing the 30 item GSRQ as well as the validated generic SF36 at home. We also asked patients some semi-structured questions about the format and content of the GSRQ. We asked patients to complete the questionnaires at baseline and four weeks.

We carried out exploratory principal components analysis on the pilot sample to determine the factorability of the data and the underlying dimensions of the GSRQ. We

examined the Bartlett's test of sphericity and the Kaiser- Meyer-Olkin (KMO) measure of sampling adequacy to determine if principal components analysis was appropriate.

Stage 3: Main study

We validated the GSRQ on a separate sample of patients. We included patients recruited to the MINuET trial to validate the GSRQ. MINuET was a 23-hospital randomised trial designed to compare gastrointestinal endoscopy (flexisigmoidoscopy and oesophagogastroduodenoscopy (OGD)) performed by doctors and nurses[22,23]. We undertook concurrent validation of the GSRQ with new patients taking part in MINuET and for whom a diagnosis had not yet been confirmed. We invited patients to complete a questionnaire containing the GSRQ, SF36 and EQ5D at recruitment, one month and 12 months. In the one and 12 month questionnaires, we asked patients if their condition had remained stable, got better or worse. We followed the practice of using self-reported global rating documented in existing literature to assess the reproducibility and responsiveness of the GSRQ[24,29,30,31,32,33]. We used patient data collected at recruitment and one month to test the psychometric properties of the GSRQ.

Patient data collection for the pilot and main study was carried out between January and April 2002.

Statistical analysis

We carried out the following psychometric analysis on the completed questionnaires from the main study:

Assessing underlying dimensions and internal consistency

We carried out principal components analysis (applying oblimin rotation). We considered that a was factor important if its eigenvalue exceeded 1. In addition we explored the face validity of the factor, that is, it appeared at "face value" to be measuring a clinically recognizable symptoms related to the patient's health. We only considered those items as contributing to a subscale if they had a factor loading of at least 0.4 on that factor[34].

We assessed the internal consistency of the GSRQ sub-scales by examining the item-total correlations (the statistical correlation of each item with the total sub-scale score) and Cronbach's alpha. We considered items for rejection if their item-total correlations were below 0.4 or above 0.8. We also examined items for floor and ceiling effects. We considered items for rejection if more than 80% of patients gave the same response because such items were not sensitive enough to discriminate different levels of severity. We examined the Cronbach's alpha for each of the resulting scales to ensure they exceeded 0.7[21,35].

Assessing validity

We evaluated the construct validity of the GSRQ questionnaire by correlation with the patients' general health as measured by the generic SF36 and EQ-5D questionnaires. If

the GSRQ and its subscales were valid measures of the effect of GI symptoms on health, we would expect that they would show significant small to moderate levels of negative correlation with the SF-36 scales and EQ5D index value[35]. In addition, if the GSRQ subscales were valid measure of GI symptoms, we would expect patients referred to receive an OGD to score worse on those GSRQ subscales related to upper GI symptoms and those referred to receive flexisigmoidoscopy to score worse on those GSRQ subscales related to lower GI symptoms.

Assessing reproducibility

We assessed reproducibility by comparing patients' GSRQ scores at recruitment and at one month. It has been suggested that a period of 1-2 weeks is the most appropriate to assess test-retest reliability[21], however we chose a period of one month to coincide with the next clinical appointment. We expected that for patients reporting no change in their gastrointestinal symptoms, their scores at recruitment and at one month should be consistent. We assessed the reproducibility of the scores for stable patients using the intraclass correlation[36].

Assessing responsiveness

We assessed the responsiveness of the GSRQ in those patients' reporting either an improvement or a deterioriation in their gastrointestinal symptoms a month after recruitment. We used Guyatt's responsiveness statistics to quantify the responsiveness of GSRQ[36].

Details of final scoring

In analysing and validating the final version of the GSRQ we calculated sub-scale scores thus:

1. As all the questions assessed response in the same direction, therefore no transformation of the scales was necessary.
2. All questions had five response options (not at all, once a week, two or three times a week, most days, every day). We scored each question in ascending order of severity from 0 to 4.
3. We gave all questions equal weighting.
4. We calculated the GSRQ sub-scale scores by summing all the responses from the final selected questions for that sub-scale and dividing by the number of valid responses.
5. We transformed the GSRQ sub-scale scores to a 0-100 scale using the formula: ((score-lowest possible score/score range) x100).
6. A higher GSRQ sub-scale score indicated worse symptom severity.

If individual question responses were missing, we still calculated sub-scale scores when at least 50% of the questions for that sub-scale had been completed. The sub-scale score was calculated by summing the responses to each answered question and dividing by the number of completed questions. If patients had completed fewer than 50% of the questions for a particular sub-scale, we treated that GSRQ sub-scale score as

missing. It was possible therefore for a patient to have scores for some of the sub-scales but not for others depending on which questions the patient had completed.

Ethics

The study was approved by the Multicentre Research Ethics Committee for Wales. We also obtained approval from all the United Kingdom participating sites local research ethics committees (Ayr Hospital; Crosshouse Hospital, Kilmarnock; Darlington Memorial Hospital; Gartnavel General Hospital, Glasgow; George Eliot Hospital, Nuneaton; Kettering General Hospital; Leicester Royal Infirmary; Royal Albert Edward Infirmary, Wigan; Monklands Hosital, Airdrie; City General Hospital, Stoke on Trent; Northampton General Hospital; Oldchurch Hospital, Romford; Queen Alexandra Hospital, Portsmouth; Queen's Medical Centre, Nottingham; Rotherham General Hospital). All patients provided written informed consent to participate in the study.

Results

Stage 1: Item generation

We developed the initial version of the GSRQ following expert review of items identified from literature (see Appendix 1). Many of the questions in the initial version were drawn from the existing questionnaires and were assimilated and modified to ensure they were appropriate to patients where no confirmed diagnosis had yet been made.

The initial questionnaire contained six sections covering a comprehensive range of common GI symptoms. Each section contained two components relating to 1) the presence of symptoms and 2) the impact of these symptoms on daily living. We designed the questionnaire to allow patients who did not have some of the symptoms in the GSRQ to skip the questions related to the impact of these symptoms on their daily living. The GSRQ took approximately 5 minutes to complete.

Stage 2: Pilot study initial validation

A total of 351 patients were recruited for the initial validation and completed the baseline questionnaire. Of these, 308 also completed the four week questionnaire. Analysis of preliminary findings from the initial validation showed three dimensions underlying GI symptoms reported by these patients. (1. Upper GI symptoms, 2. Lower GI symptoms – frequent bowel movement and related symptoms and 3. Lower GI – constipation related symptoms). Good internal consistency was recorded among the dimensions (Cronbach's alpha range 0.86 to 0.91).

Construct validity was demonstrated by statistically significant correlations between the three GSRQ dimensions with five of the eight SF36 subscales (physical functioning, role physical, pain, general health and role emotional). The upper GI dimension was also correlated with the SF36 mental health subscale. Analysis of the semi-structured questions showed that patients found the questionnaire easy to complete and there were no questions they did not wish to answer.

Stage 3: Main study

1888 patients consented to take part in MINuET, of whom 1099 were referred to received flexi-sigmoidoscopy and 789 referred to receive OGD. The questionnaire was completed by 1782 patients at recruitment and 1427 at one month.

Psychometric Analysis

Underlying dimensions

Initial exploratory principal component analysis identified a number of items that were candidates for removed from the scale. We identified that items addressing the possible impact of very different types of GI symptoms were loaded on the same factor (items 3, 9, 14, 18, 27; see Appendix 1). These were also items which patients were allowed to skip if they did not have some of the GI symptoms. Their factor loadings could therefore be reflecting a statistical artefact rather than genuine correlation. Two items (item 19- weight change and item 26- bleeding in back passage) did not contribute to any of the factors and were also considered to be candidates for removal. Item 28 (change in symptoms) similarly did not contribute to any factor and also had a comparatively high response rate to one category (70%). Item 8 (blood in vomit) also had a high response rate to one category (91%). We identified from the exploratory principal components analysis that items 29 and 30 loaded on more than one factor and formed one factor on their own. Item 29 asked patients' about their difficulty in getting to sleep and item 30 asked about the patient waking up at night. On examining the content of the two items, we felt that they were too similar and they did not make a

genuine factor. We made a decision to exclude the five optional and further six items from the final principal components analysis.

We carried principal components analysis on the remaining 19 items and identified a four-factor solution. The solution was deemed to be satisfactory and factors extracted were used in the assessment of internal consistency, construct validity, reproducibility and responsiveness. Each item was given an equal weight and was scored from 0 (not at all) to 4 (everyday). As all questions were asked in the same way, we did not carry out any transformations.

Table 1 illustrates the results of the final principal component analysis. The number of patients responding to each item ranged from 1131 to 1767, except for item 8 where 368 patients responded. We show all factor loadings of 0.4 or above, with those in parentheses illustrating that the relevant item had been excluded from one factor in favour of another.

We identified four subscales from the principal components analysis underlying the GI symptoms (see Appendix 2):

1. Upper GI symptoms – heartburn (Q1), reflux (Q2), nausea (Q3), retching (Q4), vomiting (Q5), food sticking in gullet (Q11), eating restricted (Q12) and lack of appetite (Q13).

2. Wind-related symptoms – upper abdomen discomfort (Q6), belching (Q7), wind from bowel (Q8), trapped wind (Q9) and gurgling in stomach (Q10).
3. Lower GI symptoms- frequent bowel movement (Q14), loose stools (Q15) and urgent need to empty bowel (Q18).
4. Defecation-related symptoms – hard stools (Q16), constipation (Q17) and incomplete bowel emptying (Q19).

Internal consistency

Table 2 shows the internal consistency of the four GSRQ subscales. None of the questions in the four subscales had a corrected item-subscale correlation of less than 0.40 or more than 0.8[37] . The Cronbach alphas achieved by the subscales ranged from 0.70 to 0.85 (Table 2), thus satisfying the criterion proposed by Nunnally[35] of 0.7 for comparing groups of patients.

Validity

The GSRQ demonstrated construct validity as shown by the statistically significant but modest correlations between the four GSRQ subscales with all the SF36 subscales and the EQ-5D index value (Table 3).

Results of independent t-tests which compared patients' subscale scores on the GSRQ showed that patients referred to receive an OGD scored worse on those subscales

related to upper GI symptoms and those patients referred to receive flexisigmoidoscopy scored worse on subscales related to lower GI symptoms (Table 4).

Reproducibility

Of the 1427 patients who returned both the baseline and the one month retest questionnaire, 956 (67%) did not report changes in their gastrointestinal symptoms, 371 (26%) reported improvement, 82 (6%) reported deterioration, and 18 (1%) did not give any information about the status of their symptoms. The 956 stable patients reported significantly better GSRQ scores at one month by paired t-test but the differences were small (-0.06 to -4.95). However, intraclass correlations between test and retest scores were good (0.71 to 0.77) (see Table 5).

Responsiveness

We assessed responsiveness for the 371 patients who reported an improvement and the 82 who reported deterioration in their gastrointestinal symptoms at one month. Those reporting an improvement had significantly better scores on all the four GSRQ subscales at one month by paired t-tests. Responsiveness ratios for those reporting improvements were reasonable for all the four GSRQ subscales. Those reporting a deterioration had worse GSRQ scores and the differences were significant for two of the subscales. Responsiveness ratio for those reporting deterioration were reasonable for two GSRQ subscales (see Table 6).

Discussion

The GSRQ was developed and validated in patients with luminal gastrointestinal disorders and the results illustrate that the GSRQ is a valid questionnaire for assessing symptoms in patients with different GI disorders. We were able to successfully use the questionnaire in patients where no confirmed diagnosis had been given and for patients with a confirmed disorder during the course of their disease.

We were able to develop four interpretable factors in the GSRQ relevant to upper GI, lower GI, wind-related, and defecation-related symptoms. The GSRQ had good internal consistency as demonstrated by item-total correlations and Cronbach alphas within the appropriate ranges[35,37]. Good construct validity was demonstrated by significant but modest correlations between the GSRQ sub-scales and the generic SF-36 sub-scales the and EQ5D index value. These data thereby validate that the GSRQ domains are measuring a set of gastrointestinal specific symptoms that impact on patients' overall quality of life. Variation in the GSRQ sub-domain scores is not associated with changes on generic health related quality of life instruments as different constructs are measured. Good construct validity was also demonstrated by the ability of the GSRQ to distinguish between patients with different GI conditions. Good test-retest reliability intra-class correlations and responsiveness ratios were shown for those patients reporting a change as illustrated the reproducibility and responsiveness of the GSRQ. We only used a limited patient assessment regarding whether their condition had changed. A clinical assessment of change or more extensive assessment such as the Patient Global Impression of improvement may be more appropriate.

The GSRQ questionnaire was systematically developed and piloted. Patients from 24 hospitals across the United Kingdom (one pilot and 23 main study) with a variety of luminal GI symptoms were involved in testing the questionnaire as part of a randomised controlled trial. The analysis was also thoroughly reviewed by patients, members of the public, psychometricians, statisticians and outcome specialists with experience in gastroenterology. The meticulous development and validation enhanced the robustness of GSRQ.

Given the high proportion of patients who present with luminal GI symptoms and the absence of or inappropriateness of many available tools for GI disorders, a valid and reliable system-specific scale like the GSRQ will help to monitor the long term care and disease course of the substantial amount of patients who attend GI departments with luminal gastrointestinal symptoms but have no confirmed diagnosis. The GSRQ could also provide a patient-friendly template to guide the routine electronic collection of clinical information. This would have the potential to enhance patient care at both individual and collective level.

We were not able to assess whether differences exist between those patients with or without a diagnosis. Further work is needed to determine what (if any) differences exist between these groups. In addition, further work is needed to determine whether the GSRQ can be applied to other English speaking nations and translated into non-English versions.

Main text word count (excluding abstract, tables and references)= 3656

Acknowledgments

We would like to thank Anne Seagrove and Gaynor Demery for their administrative support on the project. We would also like to thank all the clinicians for their help in recruiting patients and the patients who completed questionnaires and without whom this study would not have been possible.

References

1. Frank L, Kleinman L, Ganoczy D, McQuaid K, Sloan S, et al. (2000) Upper gastrointestinal symptoms in North America: prevalence and relationship to healthcare utilization and quality of life. Dig Dis Sci 45: 809-818.
2. Heading RC (1999) Prevalence of upper gastrointestinal symptoms in the general population: a systematic review. Scand J Gastroenterol Suppl 231: 3-8.
3. Chaplin A, Curless R, Thomson R, R. B (2000) Prevalence of lower gastrointestinal symptoms and associated consultation behaviour in a British elderly population determined by face-to-face interview. Br J Gen Pract 50: 798-802.
4. Mahmood Z, McNamara D (2003) Gastro-oesophageal reflux disease and ulcer disease. Aliment Pharmacol Ther 18 Suppl 3: 31-37.
5. Delvaux M (2003) Functional bowel disorders and irritable bowel syndrome in Europe. Aliment Pharmacol Ther 18 Suppl 3: 75-79.
6. Jones RH, Lydeard SE, Hobbs FD, Kenkre JE, Williams EI, et al. (1990) Dyspepsia in England and Scotland. Gut 31: 401-405.

7. Penston JG, Pounder RE (1996) A survey of dyspepsia in Great Britain. Aliment Pharmacol Ther 10: 83-89.

8. Stanghellini V (1999) Three-month prevalence rates of gastrointestinal symptoms and the influence of demographic factors: results from the Domestic/International Gastroenterology Surveillance Study (DIGEST). Scand J Gastroenterol Suppl 231: 20-28.

9. Harvey RF, Salih SY, Read AE (1983) Organic and functional disorders in 2000 gastroenterology outpatients. Lancet 1: 632-634.

10. Mitchell CM, Drossman DA (1987) Survey of the AGA membership relating to patients with functional gastrointestinal disorders. Gastroenterology 92: 1282-1284.

11. Wiklund I (1995) Aspects of quality of life in gastrointestinal disease: some methodological issues. Scand J Gastroenterol Suppl 208: 129-132.

12. Garratt AM, Ruta DA, Abdalla MI, Buckingham JK, Russell IT (1993) The SF36 health survey questionnaire: an outcome measure suitable for routine use within the NHS? Bmj 306: 1440-1444.

13. Dimenas E, Glise H, Hallerback B, Hernqvist H, Svedlund J, et al. (1993) Quality of life in patients with upper gastrointestinal symptoms. An improved evaluation of treatment regimens? Scand J Gastroenterol 28: 681-687.

14. Novitsky YW, Zawacki JK, Irwin RS, French CT, Hussey VM, et al. (2002) Chronic cough due to gastroesophageal reflux disease: efficacy of antireflux surgery. Surg Endosc 16: 567-571.

15. Borgaonkar MR, Irvine EJ (2000) Quality of life measurement in gastrointestinal and liver disorders. Gut 47: 444-454.

16. Rentz AM, Battista C, Trudeau E, Jones R, Robinson P, et al. (2001) Symptom and health related quality of life measures for use in selected gastrointestinal disease studies: a review and synthesis of the literature. Pharmacoeconomics 19: 349-363.

17. Yacavone RF, Locke GR, 3rd, Provenzale DT, Eisen GM (2001) Quality of life measurement in gastroenterology: what is available? Am J Gastroenterol 96: 285-297.

18. Yadegarfar G, Friend L, Jones L, Plum LM, Ardill J, et al. (2013) Validation of the EORTC QLQ-GINET21 questionnaire for assessing quality of life of patients with gastrointestinal neuroendocrine tumours. Br J Cancer 108: 301-310.

19. Drossman DA, Li Z, Andruzzi E, Temple RD, Talley NJ, et al. (1993) U.S. householder survey of functional gastrointestinal disorders. Prevalence, sociodemography, and health impact. Dig Dis Sci 38: 1569-1580.

20. Eypasch E, Williams JI, Wood-Dauphinee S, Ure BM, Schmülling C, et al. (1995) Gastrointestinal Quality of Life Index: development, validation and application of a new instrument. Br J Surg 82: 216-222.

21. Streiner GL, Norman RD (2008) Health measurement scales. A practical guide to their development and use. Oxford: Oxford University Press.

22. Richardson G, Bloor K, Williams J, Russell I, Durai D, et al. (2009) Cost effectiveness of nurse delivered endoscopy: findings from randomised multi-institution nurse endoscopy trial (MINuET). Bmj 338: b270.

23. Williams J, Russell I, Durai D, Cheung WY, Farrin A, et al. (2009) Effectiveness of nurse delivered endoscopy: findings from randomised multi-institution nurse endoscopy trial (MINuET). Bmj 338: b231.

24. Cheung WY, Garratt AM, Russell IT, Williams JG (2000) The UK IBDQ-a British version of the inflammatory bowel disease questionnaire. development and validation. J Clin Epidemiol 53: 297-306.

25. Garratt AM, Ruta DA, Russell I, Macleod K, Brunt P, et al. (1996) Developing a condition-specific measure of health for patients with dyspepsia and ulcer-related symptoms. J Clin Epidemiol 49: 565-571.

26. Velanovich V, Vallance SR, Gusz JR, Tapia FV, Harkabus MA (1996) Quality of life scale for gastroesophageal reflux disease. J Am Coll Surg 183: 217-224.

27. Hahn BA, Kirchdoerfer LJ, Fullerton S, Mayer E (1997) Evaluation of a new quality of life questionnaire for patients with irritable bowel syndrome. Aliment Pharmacol Ther 11: 547-552.

28. Svedlund J, Sjodin I, Dotevall G (1988) GSRS--a clinical rating scale for gastrointestinal symptoms in patients with irritable bowel syndrome and peptic ulcer disease. Dig Dis Sci 33: 129-134.

29. Deyo RA, Inui TS (1984) Toward clinical applications of health status measures: sensitivity of scales to clinically important changes. Health Serv Res 19: 275-289.

30. Deyo RA, Centor RM (1986) Assessing the responsiveness of functional scales to clinical change: an analogy to diagnostic test performance. J Chronic Dis 39: 897-906.

31. MacKenzie CR, Charlson ME, DiGioia D, Kelley K (1986) Can the Sickness Impact Profile measure change? An example of scale assessment. J Chronic Dis 39: 429-438.

32. Bindman AB, Keane D, Lurie N (1990) Measuring health changes among severely ill patients. The floor phenomenon. Med Care 28: 1142-1152.

33. Fitzpatrick R, Ziebland S, Jenkinson C, Mowat A (1993) Transition questions to assess outcomes in rheumatoid arthritis. Br J Rheumatol 32: 807-811.

34. Stevens JP (2002) Applied Multivariate Statistics for the Social Sciences (4th edition). Hillsdale, NJ: Erlbaum.

35. Nunnally JC, Bernstein IR (1994) Psychometric Theory (3rd Edition). New York: McGraw-Hill.

36. Deyo RA, Diehr P, Patrick DL (1991) Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. Control Clin Trials 12: 142S-158S.

37. Kline P (1986) A handbook of test construction: Introduction to psychometric design. London: Methuen.

Table 1. Selection of questions for the GSRQ and its subscales

Question No. (no of patients)	Content (Symptom)	Maximum response frequency (%)	Order of question removal	Significant factor coefficients (after rotation)			
				F1	F2	F3	F4
1 (1756)	Heartburn	51		0.45	(0.48)		
2 (1752)	Upper abdomen discomfort	44		(0.47)	0.58		
3* (1131)	Impact of symptom 1 or 2	49	1	Na	na	na	na
4 (1748)	Reflux	59		0.62			
5 (1755)	Nausea	58		0.73			
6 (1758)	Retching	79		0.76			
7 (1761)	Vomited	87		0.66			
8 (368)	Blood in vomit if vomited	91	1	Na	na	na	na
9*(1519)	Impact of symptom 4 to 7	70	1	Na	na	na	na
10 (1749)	Belch	39			0.74		
11 (1761)	Wind from bowel	28			0.58		
12 (1757)	Trapped wind	30			0.72		
13 (1759)	Gurgling in stomach	36			0.60		
14*(1748)	Impact of symptom 10 to 13	66	1	Na	na	na	na
15 (1754)	Food sticks in gullet	70		0.50			
16 (1750)	Eating restricted	56		0.54	(0.45)		
17 (1767)	Lack of appetite	64		0.62			
18* (1750)	Impact of symptom 15 to 17	73	1	Na	na	na	na
19 (1722)	Change in weight	56	2	Na	na	na	na
20 (1754)	Frequent bowel movement	47				0.84	
21 (1753)	Loose stools	43				0.85	
22 (1744)	Hard stools	58					0.82
23 (1751)	Constipation	62					0.86
24 (1753)	Urgent need to empty bowel	55				0.79	
25 (1757)	Incomplete bowel emptying	42				(0.42)	0.59
26 (1759)	Bleeding in back passage	59	2	Na	na	na	na
27*(1764)	Impact of symptom 20 to 26	65	1	Na	na	na	na
28 (1725)	Change in symptoms	72	3	Na	na	na	na
29 (1757)	Difficulty getting to sleep	58	4	Na	na	na	na

30 (1752)	Wake up	55	4	Na	na	na	na
Eigen value				5.62	2.37	1.96	1.38

* Questions which patients skip if they did not have the relevant symptoms.

Table 2. Internal consistency of the GSRQ subscales

Sub-scale	Minimum corrected item-subscale correlation	Maximum corrected item-subscale correlation	Cronbach's alpha
Upper GI	0.48	0.65	0.82
Wind-related	0.47	0.70	0.78
Lower GI	0.68	0.75	0.85
Defecation-related	0.40	0.65	0.70

Table 3. Correlation between the GSRQ subscales and the eight SF-36 subscales and EQ5D index value

GSRQ subscales SF-36 subscales and EQ5D index value	Upper GI	Wind related symptoms	Lower GI	Defecation related symptoms
I Functional status				
Physical Functioning	-0.28	-0.20	-0.14	-0.19
Social Functioning	-0.41	-0.33	-0.24	-0.26
Role limitations attributed to physical problems	-0.32	-0.26	-0.22	-0.24
Role limitations attributed to emotional problems	-0.32	-0.26	-0.22	-0.26
II Well-being				
Mental health	-0.36	-0.31	-0.22	-0.25
Vitality	-0.40	-0.37	-0.24	-0.29
Bodily pain	-0.45	-0.41	-0.22	-0.27
III Overall evaluation of health				
General health perception	-0.38	-0.33	-0.23	-0.25
EQ-5D index value	-0.40	-0.33	-0.20	-0.23

All correlations are significant at <0.001 level

Table 4. Scores for the GSRQ subscales in patients referred for different procedure types at recruitment

GSRQ subscales	Mean scores for patients referred to receive OGD	Mean scores for patients referred to receive flexisigmoidoscopy	95% CI for difference
Upper GI	25.64	13.06	(10.78, 14.37)
Wind-related	47.19	37.91	(6.80, 11.75)
Lower GI	22.88	33.87	(-13.77, -8.22)
Defecation-related	18.86	24.96	(-8.26, -3.94)

All differences significant at $p < 0.001$ level

Table 5. Reproducibility of the GSRQ subscales for stable patients

Scale	Mean difference (retest-test)	SD of the difference	95% CI of the difference	Intraclass correlation
Upper GI	-2.80	11.51	-3.54, -2.05	0.77
Wind-related	-4.95	17.59	-6.09, -3.81	0.75
Lower GI	-2.82	20.11	-4.12, -1.52	0.75
Defecation-related	-0.06	16.65	-1.14, 1.02	0.71

Table 6. Responsiveness of the GSRQ subscales for patients who improved or deteriorated

Scales	Average difference for subjects reporting a change (retest-test)	Two-tailed significance	SD of stable subjects	Responsivness ratio
Patients who improved				
Upper GI	-6.60	<0.001	11.51	0.57
Wind-related	-12.58	<0.001	17.59	0.72
Lower GI	-8.06	<0.001	20.11	0.40
Defecation-related	-4.82	<0.001	16.65	0.29
Patients who deteriorated				
Upper GI	3.92	0.005	11.51	0.34
Wind-related	1.94	0.298	17.59	0.11
Lower GI	0.15	0.956	20.11	0.008
Defecation-related	6.48	0.009	16.65	0.39

Supplementary information (Web only material)

Appendix 1: The initial version of the Gastrointestinal Symptom Rating Questionnaire (GSRQ) containing 30 items

Appendix 2: The Gastrointestinal Symptom Rating Questionnaire (GSRQ) developed following psychometric analysis (4 sub-scales; 19 items)