Article:
Ford, AC, Bercik, P, Morgan, DG et al. (3 more authors) (2013) Validation of the Rome III criteria for the diagnosis of irritable bowel syndrome in secondary care. Gastroenterology, 145 (6). 1262 - 1270. ISSN 0016-5085

https://doi.org/10.1053/j.gastro.2013.08.048

TITLE PAGE

Title: Validation of the Rome III Criteria for the Diagnosis of Irritable Bowel Syndrome in Secondary Care.

Short title: Validation of the Rome III Criteria for IBS.

Authors: Alexander C Ford[1, 2], Premysl Bercik[3], David G Morgan[4], Carolina Bolino[3], Maria Ines Pintos-Sanchez[3], Paul Moayyedi[3].

[1]Leeds Gastroenterology Institute, St. James's University Hospital, Leeds, UK.

[2]Leeds Institute of Biomedical and Clinical Sciences, University of Leeds, Leeds, UK.

[3]Farncombe Family Digestive Health Research Institute, Gastroenterology Division, McMaster University, Health Sciences Center, Hamilton, Ontario, Canada.

[4]Gastroenterology Department, St. Joseph's Healthcare, Hamilton, Ontario, Canada.

Grant support: Canadian Association of Gastroenterology (the study sponsor had no role in the study design, collection, analysis, or interpretation of data).

Abbreviations:

BMI	body mass index
CI	confidence interval
GI	gastrointestinal
IBS	irritable bowel syndrome
LR	likelihood ratio
SD	standard deviation

Correspondence: Dr. Alex Ford

Leeds Gastroenterology Institute

Room 125

4th Floor

Bexley Wing

St. James's University Hospital

Beckett Street

Leeds

United Kingdom

LS9 7TF

Email: alexf12399@yahoo.com

Telephone: +441132684963

Facsimile: +441132429722

Disclosures: ACF: none to declare. PB: none to declare. DGM: none to declare. CB: none to declare. MIP-S: none to declare. PM: none to declare.

Writing assistance: None

Author contributions: ACF, PB, DGM, CB, MIP-S, and PM conceived and drafted the study. ACF, CB, and MIP-S collected all data. ACF analyzed and interpreted the data. PM provided statistical advice and support. ACF drafted the manuscript. All authors have approved the final draft of the manuscript.

Word count: 6883

ABSTRACT

Background & aims: There are few validation studies of existing diagnostic criteria for irritable bowel syndrome (IBS). We conducted a validation study of the Rome and Manning criteria in secondary care.

Methods: We collected complete symptom, colonoscopy, and histology data from 1848 consecutive adult patients with gastrointestinal (GI) symptoms at 2 hospitals in Hamilton, Ontario; the subjects then underwent colonoscopy. Assessors were blinded to symptom status. Individuals with normal colonoscopy and histopathology examination of biopsy specimens, and no evidence of celiac disease, were classified as having no organic GI disease. The reference standard used to define the presence of true IBS was lower abdominal pain or discomfort in association with a change in bowel habit, and no organic GI disease. Sensitivity, specificity, and positive and negative likelihood ratios (LRs), with 95% confidence intervals (CIs), were calculated for each diagnostic criteria.

Results: In identifying patients with IBS, sensitivities of the criteria ranged from 61.9% (Manning) to 95.8% (Rome I), and specificities from 70.6% (Rome I) to 81.8% (Manning). Positive LRs ranged from 3.19 (Rome II) to 3.39 (Manning), and negative LRs from 0.06 (Rome I) to 0.47 (Manning). The level of agreement between diagnostic criteria was greatest for Rome I and Rome II ($\kappa = 0.95$), and lowest for Manning and Rome III ($\kappa = 0.59$).

Conclusions: Existing diagnostic criteria perform modestly in distinguishing IBS from organic disease. There appears to be little difference in terms of accuracy. More accurate ways of diagnosing IBS, avoiding the need for investigation, are required**.**

Keywords: Irritable bowel syndrome; Rome III criteria; accuracy; sensitivity; specificity

INTRODUCTION

Irritable bowel syndrome (IBS) is a functional gastrointestinal (GI) disorder, characterized by abdominal pain or discomfort, in association with altered stool form or stool frequency. [1] The condition has a relapsing and remitting natural history, [2-5] with a prevalence of up to 20% in the general population, [6] and is commoner in women. [7] Individuals with IBS are more likely to consume health care resources than healthy individuals, [8] with up to 80% of sufferers consulting their primary care physician as a result of symptoms. [9, 10]

Diagnosing IBS can be challenging due to overlap between the symptoms that sufferers report and those of organic GI conditions such as celiac disease, [11-13] small intestinal bacterial overgrowth, [14] bile acid diarrhea, [15, 16] exocrine pancreatic insufficiency, [17] or inflammatory bowel disease. [18] Partly as a result of this uncertainty, symptom-based diagnostic criteria were developed for use by physicians consulting with patients with suspected IBS as early as the 1970s, with Manning *et al.* reporting six symptoms that were commoner among individuals found ultimately to have IBS after investigation. [19] Community-based factor analysis studies demonstrate that the symptoms thought to make up IBS cluster together, lending credence to the biological plausibility of IBS as a distinct clinical entity. [20] These observations led to the development of the Rome I criteria, [21] which have been revised on two subsequent occasions to produce the Rome II and Rome III criteria. [1, 22]

Guidelines for the management of IBS from national organizations encourage physicians to make a positive diagnosis of IBS, using these symptom-based diagnostic criteria, and to avoid extensive investigation. [23-25] Accurate diagnostic criteria for IBS are of paramount importance, as they allow physicians to make the diagnosis with

confidence, hence reducing the costs of managing the condition to the health service, and reassure patients that their physician's opinion is correct. However, a recent systematic review and meta-analysis identified very few validation studies of existing symptom-based diagnostic criteria. [26] Of the available criteria, only those of Manning *et al.* had been subject to more than one validation. [19, 27-29] There was only one eligible study examining the accuracy of the Rome I criteria, [30] despite the fact that they had been published 18 years previously, and no studies of either the Rome II or Rome III criteria.

At the time this meta-analysis was published the Rome III criteria had been described only 2 years earlier. However, despite the fact that these criteria are the accepted gold-standard for reaching a diagnosis of IBS, in the intervening 5 years no validation study has been published. We have therefore conducted a study to validate the Rome III criteria, and have compared their accuracy with other available symptom-based diagnostic criteria for IBS.

METHODS

Participants and Setting

The study was conducted among all individuals newly referred from primary care to secondary care for consideration of investigation of GI symptoms. Unselected consecutive new patients aged ≥16 years were approached in the GI outpatient clinics of McMaster University Medical Center or St. Joseph's Healthcare, two hospitals in Hamilton, Ontario serving a local population of 520,000. From January 2012 to December 2012, 26% of the referrals were tertiary care in nature. There were no exclusion criteria, other than an inability to understand written English. Potentially eligible subjects were provided with a study information sheet at their initial clinic visit, prior to consultation with a Gastroenterologist. Those who agreed to participate were asked to provide written informed consent at that visit. The Hamilton Health Sciences and McMaster University research ethics board approved the study in January 2008, and recruitment continued until December 2012.

Data Collection and Synthesis

Demographic and Symptom Data

All demographic and symptom data were collected prospectively at the initial clinic visit, and hence prior to referral for colonoscopy. Basic demographic data included age, gender, ethnicity, marital status, educational level, lifestyle (tobacco and alcohol use), height (in meters), and weight (in kilograms), which were used to calculate body mass index (BMI). Symptom data were captured using the Rome III diagnostic questionnaire for the adult functional GI disorders, [31] but we also collected

data in order to examine the accuracy of the Manning, Rome I, and Rome II criteria in diagnosing IBS. All questionnaire data were entered into a database by a trained researcher who was not involved with the clinical care of the patient, thus ensuring assessors were blinded to symptom status.

Definitions of IBS

The presence or absence of Rome III-defined IBS among individual patients was assigned according to the scoring algorithm proposed for use with the Rome III questionnaire, which is detailed in Supplementary Table 1. As the questionnaire contained other symptom items, we were also able to classify the presence or absence of IBS according to the following previously accepted gold-standard symptom-based criteria, which preceded the Rome III criteria: the Manning criteria (using the presence of ≥3 of these to define IBS in our primary analysis, as this is the most widely accepted threshold, but also using ≥2 or ≥4 criteria in a sensitivity analysis), [19] the Rome I criteria, [21] and the Rome II criteria (see Supplementary Table 1). [22] We conducted further sensitivity analyses where individuals reporting lower GI alarm symptoms, including a positive family history of colorectal cancer, rectal bleeding, weight loss, or anemia were excluded from the diagnosis of IBS using any of these criteria.

In addition, in light of a recent report from a panel of experts in the field, [32] where it was felt that bloating was the most important feature of IBS, we performed three sensitivity analyses using modified versions of the Rome III criteria. In these we removed abdominal pain or discomfort from the definition and replaced it with bloating at a frequency of once per week or more, we included the presence of both abdominal pain or discomfort and bloating at a frequency of once per week or more,

Comment [AF1]: Should be four!

we increased the frequency of abdominal pain or discomfort required to meet criteria for IBS to daily, and we used abdominal pain or discomfort at a frequency of once per week or more in combination with irregular bowel habit, defined using the presence of hard stools and loose stools at a frequency of sometimes or more.

Colonoscopic and Histopathological Data

All included patients underwent complete colonoscopy to the cecum or terminal ileum, using Pentax colonoscopes (Pentax Canada, Inc), following standard bowel preparation, using either polyethylene glycol or sodium picosulphate (depending on patient and physician preference). The responsible physician performing colonoscopic examinations remained blinded to the symptom status of the patient. Findings were recorded using the endoPRO reporting system (Pentax Canada, Inc), and these reports were accessed by the study investigators in order to record the ultimate colonoscopic diagnosis for each included patient. We classified the following findings as being consistent with organic disease at colonoscopy: evidence of colitis or terminal ileitis (inflammation or ulceration), colorectal carcinoma, stricture, or evidence of radiation-induced colorectal disease. Diverticular disease, colorectal adenoma, hemorrhoids or anal fissures were not considered to represent organic disease.

Biopsy specimens were obtained at the discretion of the responsible physician performing the colonoscopy. These specimens were interpreted by experienced GI histopathologists, who remained blinded to the symptom status of the patient. Histo~~l~~~~o~~pathological findings were recorded using the MEDITECH Healthcare Reporting System (Medical Information Technology Inc, Westwood, MA), and this was accessed by the study investigators in order to record the ultimate

histopathological diagnosis. We classified the following findings as being consistent with organic disease at histopathological examination of biopsy specimens: colonic or rectal adenocarcinoma, ulcerative colitis, Crohn's disease, inflammatory bowel disease-unclassifiable, microscopic colitis, ischemic colitis, radiation enteritis, ulceration seen macroscopically at colonoscopy with non-specific inflammation on histological examination, or neuroendocrine tumour.

Definition of Organic Lower GI disease

Using these data we were able to classify patients according to the presence or absence of organic lower GI disease. Individuals with no evidence of organic disease at both colonoscopy and histopathological examination of biopsy specimens were classified as exhibiting no organic lower GI disease, while those with evidence of organic disease at either colonoscopy or histopathological examination of biopsy specimens were classified as exhibiting organic lower GI disease.

Reference Standard

The reference standard used to define the presence of true IBS was the presence of lower abdominal pain or discomfort at a frequency of at least once per week in association with a change in bowel habit, in a patient who exhibited no evidence of organic lower GI disease after colonoscopy and histological interpretation of colonic biopsies (if obtained) that would explain these symptoms, and after exclusion of celiac disease by distal duodenal biopsy obtained at upper GI endoscopy, if celiac serology was positive.

Statistical Analysis

In order to assess whether those who underwent colonoscopy were representative of all patients seen in the two GI outpatient clinics demographic data were compared between those undergoing colonoscopy who completed the symptom questionnaire, and those who completed the symptom questionnaire but did not undergo colonoscopy, using a χ^2 test for categorical data, and an independent samples *t*-test for continuous data, with a mean and standard deviation (SD). Due to multiple comparisons a 2-tailed P value of <0.01 was considered statistically significant for these analyses. We compared organic findings in those meeting the Rome III criteria for IBS, with those who did not, using Fisher's exact test, as numbers in each cell were relatively small. We measured agreement between the various diagnostic criteria for the presence of IBS using the modified Kappa statistic, where a value <0.2 indicates poor agreement and a value >0.8 indicates very good agreement beyond chance. These statistical analyses were performed using SPSS for Windows version 19.0 (SPSS Inc, Chicago, IL, USA).

The primary aim of the study was to describe the performance of the Rome III criteria for IBS in evaluating the presence of true IBS versus the reference standard. However, we also wanted to compare the performance of the Rome III criteria for IBS with previously available symptom-based diagnostic criteria including the Manning criteria, the Rome I criteria, and the Rome II criteria. The sensitivity, specificity, and positive and negative predictive values, and their 95% confidence intervals (CIs), were calculated for each of these using a Microsoft Excel spreadsheet (XP professional edition; Microsoft Corp, Redmond, WA, USA). The positive likelihood ratio (LR) and negative LR, and their 95% CIs, were also calculated using the same spreadsheet. The positive LR can be calculated from the formula: positive LR =

sensitivity / (1-specificity), while the negative LR is derived from the formula: negative LR = (1-sensitivity) / specificity. These calculations were checked using Meta-DiSc® version 1.4 (Universidad Complutense, Madrid, Spain).

RESULTS

There were a total of 4224 consecutive patients who gave informed consent and were recruited in to the study between January 2008 and December 2012 (Figure 1). The mean age of recruited subjects was 47.6 years (range 16 to 93 years) and 2617 (62.0%) were female. In total, 1981 (46.9%) of these 4224 patients underwent complete colonoscopic evaluation for their lower GI symptoms. The mean age among those undergoing colonoscopy was 49.3 years (range 16 to 90), and 1251 (63.1%) were female.

Degree of Overlap Between Different Diagnostic Criteria for IBS

There were 758 individuals who provided sufficient symptom data to examine the degree of overlap between the four different diagnostic criteria for IBS. There was agreement between all four criteria for the diagnosis of IBS in 360 (47.5%) subjects (Figure 2). Another 179 (23.6%) met the Rome I, Rome II, and Rome III criteria, but not the Manning criteria. There were 103 (13.6%) subjects who met the Manning criteria, as well as the Rome I and Rome II criteria, but not the Rome III criteria. Sixty-nine (9.1%) individuals met only the Rome I and Rome II criteria, and 11 (1.5%) only the Manning and Rome I criteria. Finally, 32 (4.2%) patients met only one set of diagnostic criteria for IBS. The level of agreement between the four symptom-based diagnostic criteria was greatest for the Rome I and Rome II criteria (Kappa statistic = 0.95) (Table 1), and lowest for the Manning and Rome III criteria (Kappa statistic = 0.59).

Validation of the Rome III Criteria for IBS

There were 1848 individuals providing complete symptom, colonoscopy, and histology data. Demographic data of all these patients, compared with the 2243 subjects who did not undergo colonoscopy, are provided in Table 2. Those undergoing colonoscopy were slightly older, of higher BMI, and were more likely to be White Caucasian, but there were no other significant differences in demographics, including the number of patients who met the Rome III criteria for IBS, between groups.

In total, 555 (30.0%) of the 1848 patients undergoing colonoscopy met the Rome III criteria for IBS. The mean age of these 555 individuals was 42.5 years, and 413 (74.4%) were female. The prevalence of organic findings after investigation in those who met the Rome III criteria for IBS, compared with those who did not, are detailed in Table 3. None of these were significantly more common among those who met criteria for IBS. Among the 365 patients with a diagnosis of IBS according to the reference standard following colonoscopy and distal duodenal biopsy (where appropriate), 251 met the Rome III criteria for IBS, giving a sensitivity of 68.8% (Table 4). Among 1483 subjects who were not judged to have IBS according to the reference standard, 1179 did not meet the Rome III criteria, giving a specificity of 79.5%. The positive LR of the Rome III criteria for the diagnosis of IBS was therefore 3.35 (95% CI 2.97 to 3.79), while the negative LR was 0.39 (95% CI 0.34 to 0.46).

Sensitivity analyses using the modifications to the Rome III criteria did not lead to any improvement in their performance (Table 4). In particular, the replacement of abdominal pain by bloating led to a decrease in the positive LR, and an increase in the negative LR, while the inclusion of both abdominal pain and bloating in the definition improved the positive LR, but the negative LR increased.

Validation of the Rome II Criteria for IBS

When we attempted to validate the Rome II criteria, there were 1864 individuals providing complete symptom, colonoscopy, and histology data. Of these, 761 (40.8%) met the Rome II criteria for IBS (mean age 43.6 years, 551 (72.4%) female). Among the 378 patients with a diagnosis of IBS according to the reference standard, 341 met the Rome II criteria for IBS, giving a sensitivity of 90.2% (Table 4). Among 1486 subjects who were not judged to have IBS according to the reference standard, 1066 did not meet the Rome II criteria, giving a specificity of 71.7%. The positive LR of the Rome II criteria for the diagnosis of IBS was therefore 3.19 (95% CI 2.92 to 3.48), while the negative LR was 0.14 (95% CI 0.10 to 0.19).

Validation of the Rome I Criteria for IBS

When we examined the validity of the Rome I criteria, there were 1866 individuals providing complete symptom, colonoscopy, and histology data. Of these, 799 (42.8%) met the Rome I criteria for IBS. The mean age of these 799 individuals was 43.8 years, and 586 (73.3%) were female. There were 377 patients with a diagnosis of IBS according to the reference standard following colonoscopy and distal duodenal biopsy (where appropriate), and 361 of these met the Rome I criteria for IBS, giving a sensitivity of 95.8% (Table 4). Among the 1489 individuals not judged to have IBS according to the reference standard, 1051 did not meet the Rome I criteria, giving a specificity of 70.6%. The positive LR of the Rome I criteria for the diagnosis of IBS was 3.26 (95% CI 3.00 to 3.53), while the negative LR was 0.06 (95% CI 0.04 to 0.10).

Validation of the Manning Criteria for IBS

Finally, when we attempted to validate the Manning criteria, there were 1854 individuals providing complete symptom, colonoscopy, and histology data. Of these, 498 (26.9%) met the Manning criteria for IBS (mean age 42.3 years, 366 (73.5%) female). Among the 367 subjects with a diagnosis of IBS according to the reference standard, 227 met the Manning criteria for IBS (sensitivity 61.9%) (Table 4). Among 1487 patients who did not have IBS according to the reference standard, 1216 did not meet the Manning criteria (specificity 81.8%). The positive LR of the Manning criteria for the diagnosis of IBS was 3.39 (95% CI 2.97 to 3.88), with a negative LR of 0.47 (95% CI 0.41 to 0.53). When ≥2 of the Manning criteria were used to define presence of IBS in a sensitivity analysis, the positive LR remained very similar (3.34; 95% CI 3.04 to 3.68), but the negative LR fell (0.20; 95% CI 0.16 to 0.26). However, when ≥4 criteria were used, again the positive LR remained similar (3.42; 95% CI 2.80 to 4.18), while the negative LR increased to 0.71 (95% CI 0.66 to 0.77).

Sensitivity Analysis Excluding Individuals Reporting Lower GI Alarm Symptoms from the Definition of IBS

When those with one or more lower GI alarm symptoms were excluded from the definition of IBS, the specificities of all the different diagnostic criteria increased to >90% in all cases, but this came at the expense of sensitivity which fell dramatically (range 13.7% to 24.3%) (Table 5). As a result, positive LRs increased slightly, but negative LRs increased dramatically, meaning that the absence of any of the various diagnostic criteria performed poorly in ruling out IBS.

DISCUSSION

This study has attempted to validate the Rome III criteria for IBS against an accepted reference standard, and to compare their performance against all previous symptom-based diagnostic criteria. It has demonstrated that the presence of the Rome III criteria in a patient with lower GI symptoms increases the likelihood of having IBS more than three-fold, whilst their absence reduces the likelihood of IBS by 60%. These LRs mean that the Rome III criteria performed only modestly in predicting a diagnosis of IBS, although the application of other existing symptom-based diagnostic criteria yielded broadly similar LRs, with the exception of the negative LR for the Rome I criteria, which was 0.06, meaning that the absence of the Rome I criteria reduces the likelihood of a person having IBS by 94%. When the absence of lower GI alarm symptoms was incorporated into the various diagnostic criteria in a sensitivity analysis, their performance was even less optimal. Almost 50% of patients with IBS met all four symptom-based diagnostic criteria, and one in four individuals met all three iterations of the Rome criteria. Agreement between the various diagnostic criteria was good to excellent in most cases.

Strengths of this study include the large sample size, with over 1800 individuals undergoing colonoscopy and providing complete symptom data, meaning that this is one of the largest studies to validate available symptom-based diagnostic criteria for IBS ever conducted. It is also, to our knowledge, the first study to validate all existing diagnostic criteria for IBS simultaneously. We also performed sensitivity analyses using the available diagnostic criteria for IBS. In addition, the study was designed to adhere closely to the STARD guidelines for the reporting of studies of diagnostic accuracy, with consecutive patients recruited, assessors blinded, and an

accepted reference standard used. Finally, the fact that the majority of patients we recruited were unselected referrals to secondary care means that the results are likely to be generalizable to Gastroenterologists consulting with individuals with suspected IBS in usual clinical practice.

Weaknesses of the study include the fact that we did not mandate colonoscopy in all individuals with lower GI symptoms as part of the study design. This means that in some individuals with IBS the diagnosis will have been made on clinical grounds alone, and these individuals will not have been subject to colonoscopy. If the diagnosis of IBS were correct in all these patients then the true positive rate of all the symptom-based diagnostic criteria we assessed will have been artificially reduced, leading to an underestimation of their accuracy. In addition, those who did undergo colonoscopy and provide complete symptom data were not entirely representative of the entire study population, with an over representation of White Caucasians, older individuals, and patients with a higher BMI. However, in most cases the absolute differences in demographic data between those undergoing colonoscopy and providing complete symptom data and those who did not undergo colonoscopy were modest. We did not perform longitudinal follow-up to ensure that an organic diagnosis was not missed in individuals who met the reference standard for IBS, although previous studies have suggested that a diagnosis of IBS is unlikely to be revised during extended follow-up, [33] despite repeated investigation. [34] The reference standard included symptom data from the questionnaire, which may have led to an overestimation of the accuracy of the Rome III criteria. Finally, we did not attempt to validate a new set of diagnostic criteria for IBS as part of this study, although this was not the primary aim.

As this study was conducted within usual clinical practice, and there is no accepted gold-standard for the diagnosis of IBS, we did not mandate a minimum level of blood work, such as complete blood count, C-reactive protein, or celiac serology, in all individuals, nor did we make any recommendations as to whether upper GI or small bowel investigations should be performed. Our study assumed that where initial blood tests were abnormal, these would have prompted the responsible physician to request further appropriate investigations to exclude organic disease. However, where celiac serology was positive, distal duodenal biopsy was performed, and those individuals with celiac disease were classified as having organic disease within our analyses. The relevance of these issues is debatable, as a recent study suggests that, within a North American population, there is no difference in the prevalence of celiac disease in non-constipated IBS patients compared with controls, [35] and previous studies that have applied a routine panel of blood tests or small bowel investigations in patients with suspected IBS demonstrate a yield for organic disease of ≤1%. [13, 36, 37]

In a previous meta-analysis of four validation studies of the Manning criteria for IBS, [19, 27-29] containing a total of 574 patients, the positive and negative LRs of the Manning criteria were 2.9 and 0.29 respectively. [26] The Manning criteria performed best in the original validation study, [19] and when this study was excluded from the analysis the positive and negative LRs fell to 2.7 and 0.33. The Rome I criteria, which were validated in only one study containing 602 patients, [30] yielded a positive LR of 4.8 and a negative LR of 0.34. [26] Several of the studies included in this meta-analysis classified subjects with diverticular disease, colorectal adenomas, or both, as having organic disease. This may have led to the misclassification of individuals with true IBS as having organic disease, a problem that does not apply to our study, as we were careful to exclude these individuals from our organic lower GI disease category.

A systematic review published in 2012 once again failed to identify any studies validating the Rome III criteria. [38] More recently, a study from Denmark has attempted to examine the accuracy of the Rome III criteria for IBS, [39] compared with a reference standard of a primary care physician's diagnosis of IBS. Among 499 patients with suspected IBS the Rome III criteria were fulfilled by 376, yielding a sensitivity of 75%, similar to that we observed. However, as only patients with suspected IBS were recruited, the specificity and likelihood ratios for the Rome III criteria could not be estimated in this study.

The degree of agreement between the various diagnostic criteria was good to very good in all instances. This is probably not unexpected, as the various criteria are derived from each other, meaning that the same strengths and weaknesses are, for the most part, passed on from one set of criteria to another. It was lowest for the Manning and Rome III criteria. This is in keeping with a Greek study based in primary care, [40] and may reflect the number of revisions that the Rome III criteria have undergone. Our observation that the Rome I and II criteria demonstrated high levels of agreement is similar to that of Boyce *et al.* who reported that over 90% of individuals meeting Rome I criteria also met Rome II criteria, [41] although other investigators have shown that the level of agreement between the Rome I and Rome II criteria is lower. [42, 43] Previous studies have also suggested that agreement between the Rome I and Manning criteria are poor, [41, 44] leading some to propose that the Rome criteria are too restrictive. [41]

Our findings show that, regardless of attempts to tighten the definition of IBS by refining the symptoms used to diagnose the condition, there is little to choose between the available diagnostic criteria. Despite a high prevalence of IBS in the population under study, the positive predictive value of all these criteria was less than

50%, suggesting that more accurate ways of detecting IBS are required if the condition is to be diagnosed with any certainty, and the potential for unnecessary investigation avoided. The Rome III criteria appeared to perform less well than previous iterations, perhaps implying that they should be reappraised and brought more in line with earlier definitions. In fact, a recent survey of internationally-recognized experts reported that 80% of respondents felt that the Rome III criteria did not reflect IBS adequately, in terms of the patients they saw in clinical practice, and an identical proportion felt that a new international standardized set of criteria for the diagnosis of IBS were required. [32]

Even the exclusion of individuals reporting one or more lower GI alarm symptoms from the various diagnostic criteria for IBS did not improve their performance. This suggests that the presence of alarm symptoms in an individual does not predict organic disease with any great accuracy, an issue we have identified in a previous meta-analysis of observational studies. [45] Biomarkers may hold some promise, [46, 47] although in the only fully published study to assess these they were no more accurate than existing diagnostic criteria. [48] In the absence of an accurate and accepted biomarker for the condition, and given that our sensitivity analyses using the Rome III criteria did not improve their performance, it would seem that further refinement of the symptoms that are used to define IBS is unlikely to prove fruitful.

The lack of a gold standard for the diagnosis of a clinical entity is not unique to Gastroenterology and we need to use approaches that other disciplines, such as Psychiatry, have used to overcome this problem. One possibility is to also evaluate patients with colonoscopy and relevant biological markers, as well as obtain a detailed symptom history. [47, 49, 50] The symptom data can then be compared with the biological markers and colonoscopy results using statistical techniques, such as latent class

analysis and Bayesian analysis. [51] These techniques can classify patients into categories based on the underlying structure of the data, without relying on a gold standard. Whilst these approaches have limitations, they are recognized as being a significant methodological advance, and are recommended when there is no reference standard. [52] Our data suggest the Gastroenterology research community should take note and use new approaches to refine diagnostic criteria for IBS.

ACKNOWLEDGEMENTS

We are grateful to June Urquhart for entry of questionnaire data and administering questionnaires to patients attending clinic, and Sandra Arthur for administering questionnaires to patients attending clinic.

REFERENCES

1. Longstreth GF, Thompson WG, Chey WD, *et al*. Functional bowel disorders. Gastroenterology 2006;130:1480-1491.

2. Agreus L, Svardsudd K, Nyren O, *et al*. Irritable bowel syndrome and dyspepsia in the general population: Overlap and lack of stability over time. Gastroenterology 1995;109:671-680.

3. Agreus L, Svardsudd K, Talley NJ, J *et al*. Natural history of gastroesophageal reflux disease and functional abdominal disorders. Am J Gastroenterol 2001;96:2905-2914.

4. Ford AC, Forman D, Bailey AG, *et al*. Irritable bowel syndrome: A 10-year natural history of symptoms, and factors that influence consultation behavior. Am J Gastroenterol 2008;103:1229-1239.

5. Halder SLS, Locke III GR, Schleck CD, *et al*. Natural history of functional gastrointestinal disorders: A 12-year longitudinal population-based study. Gastroenterology 2007;133:799-807.

6. Lovell RM, Ford AC. Global prevalence of, and risk factors for, irritable bowel syndrome: A meta-analysis. Clin Gastroenterol Hepatol 2012;10:712-721.

7. Lovell RM, Ford AC. Effect of gender on prevalence of irritable bowel syndrome in the community: Systematic review and meta-analysis. Am J Gastroenterol 2012;107:991-1000.

8. Talley NJ, Gabriel SE, Harmsen WS, *et al*. Medical costs in community subjects with irritable bowel syndrome. Gastroenterology 1995;109:1736-1741.

9. Koloski NA, Talley NJ, Boyce PM. Epidemiology and health care seeking in the functional GI disorders: A population-based study. Am J Gastroenterol 2002;97:2290-2299.

10. Koloski NA, Talley NJ, Huskic SS, *et al*. Predictors of conventional and alternative health care seeking for irritable bowel syndrome and functional dyspepsia. Aliment Pharmacol Ther 2003;17:841-851.

11. Ford AC, Chey WD, Talley NJ, *et al.* Yield of diagnostic tests for celiac disease in subjects with symptoms suggestive of irritable bowel syndrome: Systematic review and meta-analysis. Arch Intern Med 2009;169:651-658.

12. Sainsbury A, Sanders DS, Ford AC. Prevalence of irritable bowel syndrome-type symptoms in patients with celia disease: A meta-analysis. Clin Gastroenterol Hepatol 2012;11:359-365.

13. Sanders DS, Carter MJ, Hurlstone DP, *et al*. Association of adult coeliac disease with irritable bowel syndrome: A case-control study in patients

fulfilling ROME II criteria referred to secondary care. Lancet 2001;358:1504-1508.

14. Ford AC, Spiegel BMR, Talley NJ, *et al*. Small intestinal bacterial overgrowth in irritable bowel syndrome: Systematic review and meta-analysis. Clin Gastroenterol Hepatol 2009;7:1279-1286.

15. Gracie DJ, Kane JS, Mumtaz S, *et al*. Prevalence of, and predictors of, bile acid malabsorption in outpatients with chronic diarrhea. Neurogastroenterol Motil 2012;24:983-e538.

16. Wedlake L, A'Hern R, Russell D, *et al*. Systematic review: The prevalence of idiopathic bile acid malabsorption as diagnosed by SeHCAT scanning in patients with diarrhoea-predominant irritable bowel syndrome. Aliment Pharmacol Ther 2009;30:707-717.

17. Leeds JS, Hopper AD, Sidhu R, *et al*. Some patients with irritable bowel syndrome may have exocrine pancreatic insufficiency. Clin Gastroenterol Hepatol 2010;8:433-438.

18. Halpin SJ, Ford AC. Prevalence of symptoms meeting criteria for irritable bowel syndrome in inflammatory bowel disease: Systematic review and meta-analysis. Am J Gastroenterol 2012;107:1474-1482.

19. Manning AP, Thompson WG, Heaton KW, *et al*. Towards positive diagnosis of the irritable bowel. Br Med J 1978;277:653-654.

20. Whitehead WE, Crowell MD, Bosmajian L, *et al*. Existence of irritable bowel syndrome supported by factor analysis of symptoms in two community samples. Gastroenterology 1990;98:336-340.

21. Drossman DA, Thompson WG, Talley NJ. Identification of sub-groups of functional gastrointestinal disorders. Gastroenterology Intl 1990;3:159-72.

22. Thompson WG, Longstreth GF, Drossman DA, *et al*. Functional bowel disorders and functional abdominal pain. Gut 1999;45 (suppl II):II43-II47.

23. American College of Gastroenterology IBS Task Force. An evidence-based position statement on the management of irritable bowel syndrome. Am J Gastroenterol 2009;104 (suppl I):S1-S7.

24. Spiller R, Aziz Q, Creed FEA, *et al*. Guidelines on the irritable bowel syndrome: Mechanisms and practical management. Gut 2007;56:1770-1798.

25. Drossman DA, Camilleri M, Mayer EA, *et al*. AGA technical review on irritable bowel syndrome. Gastroenterology 2002;123:2108-2131.

26. Ford AC, Talley NJ, Veldhuyzen Van Zanten SJ, *et al*. Will the history and physical examination help establish that irritable bowel syndrome is causing this patient's lower gastrointestinal tract symptoms? JAMA 2008;300:1793-1805.

27. Dogan UB, Unal S. Kruis scoring system and Manning's criteria in diagnosis of irritable bowel syndrome: Is it better to use combined? Acta Gastroenterol Belg 1996;59:225-228.

28. Jeong H, Lee HR, Yoo BC, *et al*. Manning criteria in irritable bowel syndrome: Its diagnostic significance. Korean J Intern Med 1993;8:34-39.

29. Rao KP, Gupta S, Jain AK, *et al*. Evaluation of Manning's criteria in the diagnosis of irritable bowel syndrome. J Assoc Physicians India 1993;41:357-363.

30. Tibble JA, Sigthorsson G, Foster R, *et al*. Use of surrogate markers of inflammation and Rome criteria to distinguish organic from nonorganic intestinal disease. Gastroenterology 2002;123:450-460.

31. Whitehead WE, and the Validation Working Team Committee in association with the Rome Questionnaire Committee. Development and validation of the Rome III diagnostic questionnaire. In: Drossman DA, editor Rome III: The functional gastrointestinal disorders, 3rd edition Virginia: Degnon Associates Inc 2006;835-853.

32. Pimentel M, Talley NJ, Quigley EM, *et al*. Report from the multinational irritable bowel syndrome initiative 2012. Gastroenterology 2013;144:e1-5.

33. Owens DM, Nelson DK, Talley NJ. The irritable bowel syndrome: Long-term prognosis and the physician-patient interaction. Ann Intern Med 1995;122:107-112.

34. Adeniji OA, Barnett CB, Di Palma JA. Durability of the diagnosis of irritable bowel syndrome based on clinical criteria. Dig Dis Sci 2004;49:572-574.

35. Cash BD, Rubenstein JH, Young PE, *et al*. The prevalence of celiac disease among patients with nonconstipated irritable bowel syndrome is similar to controls. Am J Gastroenterol 2011;141:1187-1193.

36. Banerjee R, Choung OW, Gupta R, T *et al*. Rome I criteria are more sensitive than Rome II for diagnosis of irritable bowel syndrome in Indian patients. Indian J Gastroenterol 2005;24:164-166.

37. Tolliver BA, Herrera JL, DiPalma JA. Evaluation of patients who meet clinical criteria for irritable bowel syndrome. Am J Gastroenterol 1994;89:176-178.

38. Dang J, Ardila-Hani A, Amichai MM, *et al*. Systematic review of diagnostic criteria for IBS demonstrates poor validity and utilization of Rome III. Neurogastroenterol Motil 2012;24:853-e397.

39. Engsbro AL, Begtrup LM, Kjeldsen J, *et al*. Patients suspected of irritable bowel syndrome - Cross-sectional study exploring the sensitivity of Rome III criteria in primary care. Am J Gastroenterol 2013;108:972-980.

40. Anastasiou F, Mouzas IA, Moschandreas J, *et al*. Exploring the agreement between diagnostic criteria for IBS in primary care in Greece. BMC Res Notes 2008;1:127.

41. Boyce PM, Koloski NA, Talley NJ. Irritable bowel syndrome according to varying diagnostic criteria: Are the new Rome II criteria unnecessarily restrictive for research and practice? Am J Gastroenterol 2000;95:3176-3183.

42. Mearin F, Badia X, Balboa A, *et al*. Irritable bowel syndrome prevalence varies enormously depending on the employed diagnostic criteria: Comparison of Rome II versus previous criteria in a general population. Scand J Gastroenterol 2001;36:1155-1161.

43. Mearin F, Roset M, Badia X, *et al*. Splitting irritable bowel syndrome: From original Rome to Rome II criteria. Am J Gastroenterol 2004;99:122-130.

44. Saito YA, Locke GR, Talley NJ, *et al*. A comparison of the Rome and Manning criteria for case identification in epidemiological investigations of irritable bowel syndrome. Am J Gastroenterol 2000;95:2816-2824.

45. Ford AC, Veldhuyzen Van Zanten SJO, Rodgers CC, *et al*. Diagnostic utility of alarm features for colorectal cancer: Systematic review and meta-analysis. Gut 2008;57:1545-1553.

46. Jones MP, Chey WD, Gong H, *et al*. Psychological variables add incremental value to biological markers in differentiating IBS from healthy volunteers. Gastroenterology 2012;142 (suppl 1):S820.

47. Lembo AJ, Neri B, Tolley J, *et al*. Use of serum biomarkers in a diagnostic test for irritable bowel syndrome. Aliment Pharmacol Ther 2009;29:834-842.

48. Ford AC, Talley NJ, Moayyedi P. 10-biomarker algorithm to identify irritable bowel syndrome. Aliment Pharmacol Ther 2009;30:95-96.

49. Chang L, Adeyemo M, Karagiannides I, *et al*. Serum and colonic mucosal immune markers in irritable bowel syndrome. Am J Gastroenterol 2012;107:262-272.

50. Schmulson M, Chey WD. Abnormal immune regulation and low-grade inflammation in IBS: Does one size fit all? Am J Gastroenterol 2012;107:273-275.

51. Moayyedi P, Duffy J, Delaney B. New approaches to enhance the accuracy of the diagnosis of reflux disease. Gut 2004;53 (suppl 4):iv55-iv57.

52. Rutjes AW, Reitsma JB, Coomarasamy A, *et al*. Evaluation of diagnostic tests when there is no gold standard. A review of methods. Health Technol Assess 2007;11:iii, ix-51.

Table 1. Kappa Statistic for Levels of Agreement Between the Rome and Manning Criteria for Irritable Bowel Syndrome.

	Rome III criteria	**Rome II criteria**	**Rome I criteria**	**Manning criteria (≥3 criteria)**
Rome III criteria		0.79	0.74	0.59
Rome II criteria	0.79		0.95	0.67
Rome I criteria	0.74	0.95		0.66
Manning criteria (≥3 criteria)	0.59	0.67	0.66	

Table 2. Demographics and Baseline Characteristics of Patients Undergoing Colonoscopy and Providing Complete Rome III Symptom Data, Compared with Those Who Did Not Undergo Colonoscopy.

	Underwent colonoscopy and provided complete Rome III symptom data (n = 1848)	**Did not undergo colonoscopy (n = 2243)**	**P value***
Mean age (SD)	48.9 (17.1)	46.1 (18.1)	<0.001
Mean body mass index (SD)	27.3 (6.0)	26.7 (6.25)	0.009
Female gender (%)	1185 (64.1)	1366 (60.9)	0.03
Tobacco user (%)	375 (20.3)	442 (19.7)	0.76
Alcohol user (%)	1106 (59.8)	1250 (55.7)	0.02
Marital status (%)			
Married or co-habiting	1130 (61.1)	1280 (57.1)	0.04
Divorced or separated	211 (11.4)	249 (11.1)	
Never married	406 (22.0)	568 (25.3)	
Widowed	82 (4.4)	102 (4.5)	

Educational level (%)			
Elementary	85 (4.6)	87 (3.9)	0.31
High school	520 (28.1)	627 (28.0)	
College or technical school	549 (29.7)	648 (28.9)	
University	493 (26.7)	546 (24.3)	
Postgraduate	170 (9.2)	249 (11.1)	
Ethnicity (%)			
White Caucasian	1678 (90.8)	1917 (85.5)	0.004
South Asian	19 (1.0)	41 (1.8)	
Middle-Eastern	19 (1.0)	38 (1.7)	
First Nations	18 (1.0)	18 (0.8)	
African	20 (1.1)	31 (1.4)	
South-East Asian	13 (0.7)	35 (1.6)	
Latin-American	13 (0.7)	22 (1.0)	
Met Rome III criteria for IBS (%)	555 (30.0)	579 (25.8)	0.19

*P value for independent samples *t*-test for continuous data and Pearson χ^2 for comparison of categorical data.

Table 3. Prevalence of Organic Disease in Patients Meeting the Rome III Criteria for IBS, Compared with Those Who Did Not.

	Met Rome III criteria for IBS (n = 555)	**Did not meet Rome III criteria for IBS (n = 1293)**	**P value***
Crohn's disease (%)	48 (8.6)	84 (6.5)	0.11
Ulcerative colitis (%)	34 (6.1)	59 (4.6)	0.16
Indeterminate colitis (%)	24 (4.3)	42 (3.2)	0.27
Colorectal cancer (%)	13 (2.3)	28 (2.2)	0.86
Lymphocytic colitis (%)	8 (1.4)	15 (1.2)	0.65
Celiac disease (%)	8 (1.4)	13 (1.0)	0.47
Radiation enteritis (%)	8 (1.4)	9 (0.7)	0.18
Collagenous colitis (%)	4 (0.7)	4 (0.3)	0.25
Non-specific GI ulceration (%)	3 (0.5)	2 (0.2)	0.16

*P value for Fisher's exact test for comparison of categorical data.

Table 4. Sensitivity, Specificity, Positive and Negative Predictive Values, and Positive and Negative Likelihood Ratios for the Rome and Manning Criteria for Irritable Bowel Syndrome.

	Sensitivity (95% CI)	**Specificity (95% CI)**	**Positive predictive value (95% CI)**	**Negative predictive value (95% CI)**	**Positive likelihood ratio (95% CI)**	**Negative likelihood ratio (95% CI)**
Rome III criteria	68.8% (63.8% – 73.3%)	79.5% (77.4% – 81.5%)	45.2% (41.1% – 49.4%)	91.2% (89.5% – 92.6%)	3.35 (2.97 – 3.79)	0.39 (0.34 – 0.46)
Rome II criteria	90.2% (86.8% – 92.8%)	71.7% (69.4% – 74.0%)	44.8% (41.3% – 48.4%)	96.6% (95.4% – 97.6%)	3.19 (2.92 – 3.48)	0.14 (0.10 – 0.19)
Rome I criteria	95.8% (93.2% – 97.4%)	70.6% (68.2% – 72.8%)	45.2% (41.8% – 48.7%)	98.5% (97.6% – 99.1%)	3.26 (3.00 – 3.53)	0.06 (0.04 – 0.10)
Manning criteria (≥2 criteria)	85.0% (81.0% – 88.2%)	74.6% (72.3% – 76.7%)	45.8% (42.2% – 49.5%)	95.1% (93.8% – 96.2%)	3.34 (3.04 – 3.68)	0.20 (0.16 – 0.26)

Manning criteria (≥3 criteria)	61.9% (56.8% – 66.7%)	81.8% (79.7% – 83.7%)	45.6% (41.3% – 50.0%)	89.7% (88.0% – 91.2%)	3.39 (2.97 – 3.88)	0.47 (0.41 – 0.53)
Manning criteria (≥4 criteria)	36.1% (31.3% – 41.1%)	89.5% (87.8% – 90.9%)	45.4% (39.7% – 51.1%)	85.2% (83.4% – 86.9%)	3.42 (2.80 – 4.18)	0.71 (0.66 – 0.77)
Rome III criteria with abdominal pain or discomfort replaced by bloating	54.3% (49.2% – 59.4%)	76.4% (74.2% – 78.5%)	36.9% (32.9% – 41.0%)	86.8% (84.9% – 88.6%)	2.31 (2.02 – 2.63)	0.60 (0.53 – 0.67)

Rome III criteria with abdominal pain or discomfort and bloating	53.1% (47.9% – 58.3%)	85.1% (83.2% – 86.9%)	46.7% (41.8% – 51.6%)	88.1% (86.4% – 89.8%)	3.58 (3.06 – 4.17)	0.55 (0.49 – 0.61)
Rome III criteria with daily abdominal pain or discomfort	29.0% (24.7% – 33.7%)	92.0% (90.5% – 93.2%)	48.1% (41.7% – 54.5%)	83.5% (81.6% – 85.2%)	3.61 (2.87 – 4.55)	0.77 (0.72 – 0.82)
Rome III criteria with irregular bowel habit	34.4% (29.8% – 39.3%)	91.1% (89.6% – 92.5%)	49.1% (43.1% – 55.1%)	84.9% (83.0% – 86.5%)	3.88 (3.14 – 4.81)	0.72 (0.67 – 0.78)

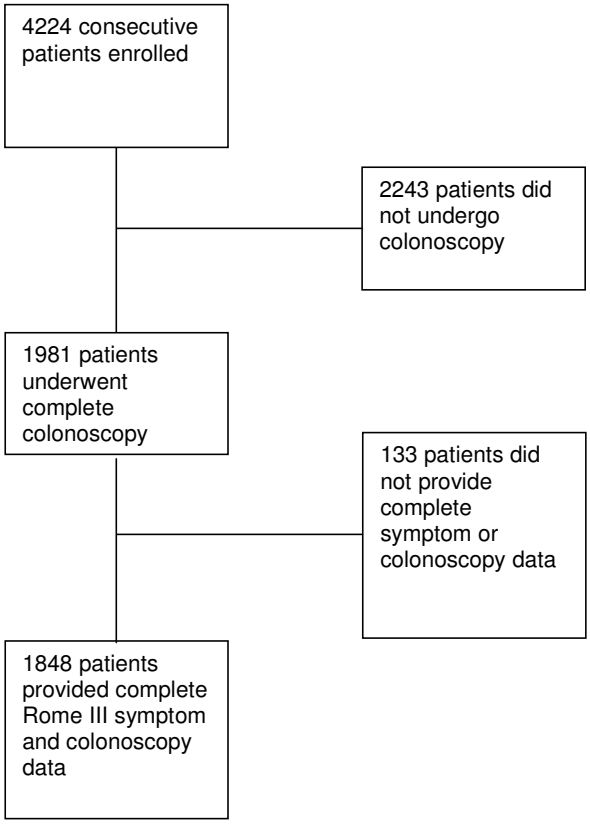
Table 5. Sensitivity, Specificity, Positive and Negative Predictive Values, and Positive and Negative Likelihood Ratios for the Rome and Manning Criteria for Irritable Bowel Syndrome, Excluding Individuals Reporting Lower GI Alarm Symptoms from the Definition of IBS.

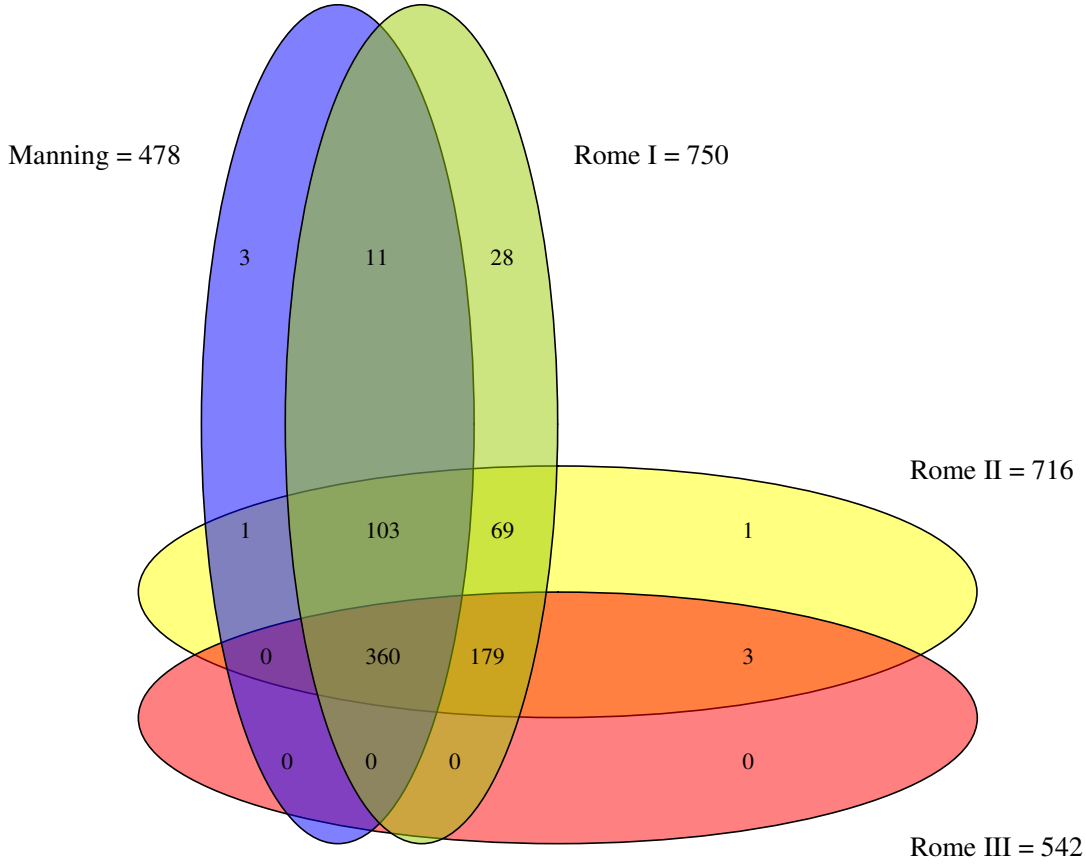
	Sensitivity (95% CI)	**Specificity (95% CI)**	**Positive predictive value (95% CI)**	**Negative predictive value (95% CI)**	**Positive likelihood ratio (95% CI)**	**Negative likelihood ratio (95% CI)**
Rome III criteria	17.4% (13.9% – 21.5%)	95.6% (94.4% – 96.5%)	49.6% (42.0% – 58.7%)	82.1% (80.0% – 83.6%)	3.92 (2.85 – 5.38)	0.86 (0.83 – 0.91)
Rome II criteria	23.3% (19.4% – 27.8%)	94.5% (93.2% – 95.5%)	51.7% (44.9% – 59.5%)	82.9% (80.8% – 84.4%)	4.21 (3.20 – 5.53)	0.81 (0.77 – 0.86)
Rome I criteria	24.3% (20.3% – 28.8%)	93.9% (92.6% – 95.0%)	50.5% (44.0% – 58.1%)	83.0% (80.9% – 84.4%)	4.01 (3.08 – 5.22)	0.81 (0.76 – 0.85)
Manning criteria (≥3 criteria)	13.7% (10.6% – 17.6%)	97.1% (96.1% – 97.8%)	54.1% (45.3% – 64.6%)	81.6% (79.6% – 83.1%)	4.66 (3.18 – 6.82)	0.89 (0.85 – 0.93)

FIGURE LEGENDS

Figure 1. Flow of Study Participants.

Figure 2. Overlap Between Diagnostic Criteria for Irritable Bowel Syndrome.


4224 consecutive patients enrolled
2243 patients did not undergo colonoscopy
1981 patients underwent complete colonoscopy
133 patients did not provide complete symptom or colonoscopy data
1848 patients provided complete Rome III symptom and colonoscopy data


Manning = 478
Rome I = 750
Rome II = 716
Rome III = 542
3
11
28
1
103
69
1
0
360
179
3
0
0
0
0