Evidence- and Consensus-Based Practice Guidelines for the Diagnosis of Irritable Bowel Syndrome

Ronnie Fass, MD; George F. Longstreth, MD; Mark Pimentel, MD; Steven Fullerton, MPH; Simcha M. Russak, MA; Chiu-Fang Chiu, PhD; Eileen Reyes, BS; Paul Crane, MD; Glenn Eisen, MD; Bill McCarberg, MD; Joshua Ofman, MD, MSHS

Background: Irritable bowel syndrome (IBS) presents a significant diagnostic and management challenge for primary care practitioners. Improving the accuracy and timeliness of diagnosis may result in improved quality and efficiency of care.

Objective: To systematically appraise the existing diagnostic criteria and combine the evidence with expert opinion to derive evidence- and consensus-based guidelines for a diagnostic approach to patients with suspected IBS.

Methods: We performed a systematic literature review (January 1966–April 2000) of computerized bibliographic databases. Articles meeting explicit inclusion criteria for diagnostic studies in IBS were subjected to critical appraisal, which formed the basis of guideline statements presented to an expert panel. To develop a diagnostic algorithm, an expert panel of specialists and primary care physicians was used to fill in gaps in the literature. Consensus was developed using a modified Delphi technique.

Results: The systematic literature review identified only 13 published studies regarding the effectiveness of competing diagnostic approaches for IBS, the accuracy of diagnostic tests, and the internal validity of current diagnostic symptom criteria. Few studies met accepted methodological criteria. While symptom criteria have been validated, the utility of endoscopic and other diagnostic interventions remains unknown. An analysis of the literature, combined with consensus from experienced clinicians, resulted in the development of a diagnostic algorithm relevant to primary care that emphasizes a symptom-based diagnostic approach, refers patients with alarm symptoms to subspecialists, and reserves radiographic, endoscopic, and other tests for referral cases. The resulting algorithm highlights the reliance on symptom criteria and comprises a primary module, 3 submodules based on the predominant symptom pattern (constipation, diarrhea, and pain) and severity level, and a subspecialist referral module.

Conclusions: The dearth of available evidence highlights the need for more rigorous scientific validation to identify the most accurate methods of diagnosing IBS. Until such time, the diagnostic algorithm presented herein could inform decision making for a range of providers caring for primary care patients with abdominal discomfort or pain and altered bowel function suggestive of IBS.

Arch Intern Med. 2001;161:2081-2088

IRRITABLE BOWEL syndrome (IBS) is a common disorder characterized by abdominal pain, bloating, and disturbed defecation. Irritable bowel syndrome remains the most common disorder encountered by gastroenterologists.1 The incidence of IBS is reported to be 15% to 20% in the general population,2,3 with prevalence rates dependent upon the symptom criteria4 used to define the condition. Furthermore, functional bowel complaints such as IBS are responsible for nearly 50% of visits to gastroenterologists.5

Presently there are no known biochemical or structural markers for identifying patients with IBS. In most cases, a diagnosis of IBS is based on typical symptoms and negative results of a limited diagnostic evaluation. Consequently, symptom criteria for diagnosis have been proposed. Currently, the most widely accepted criteria include the Rome I criteria,6 the Manning criteria,7 and the recently developed Rome II consensus criteria.8 These criteria have been used in research protocols to facilitate study inclusion. However, they have undergone limited validation, particularly in primary care settings.

In addition to symptom criteria, several diagnostic algorithms, such as that proposed by Schmulson and Chang,9 have been developed to facilitate the diagnosis and management of IBS. However, most guidelines were developed for use in the specialty care setting and targeted for patients with more severe symptoms. While many algorithms use structured and validated expert panel methods, most panels consisted of academic center subspecialists, who may not reflect the understanding and concerns of provider organizations focused on the provision of care in the primary care setting.
The objective of the present study was to arrive at evidence- and consensus-based guidelines for a diagnostic approach to patients with suspected IBS with specific relevance to primary care providers. Our evidence-based approach to guideline development relied upon a systematic review of the published medical literature and consensus from an expert panel of experienced clinicians in a variety of health care settings. A modified Delphi technique was used for instances in which the published literature could not inform the decision-making process. The purpose of this effort is to provide primary care and nonacademic clinicians with guidance to improve the evaluation and diagnosis of patients with IBS.

**METHODS**

A 3-phase approach was used to construct the evidence-based guidelines for diagnosing IBS. The phases include a systematic literature review of diagnostic studies in IBS, a comprehensive appraisal of prior studies and estimates of the accuracy of diagnostic tests, and convening an expert panel to synthesize this information and develop consensus-based recommendations about the diagnosis of IBS.

**SYSTEMATIC REVIEW OF THE LITERATURE**

We searched 4 computerized bibliographic databases (MEDLINE, HEALTHSTAR, Evidence-Based Medicine, and the Cochrane Database) to identify English-language articles published between January 1966 and April 2000. The focus of the search was on articles that evaluated the performance of diagnostic tests and procedures for IBS. Search terms and strategies were developed in cooperation with an expert librarian experienced in advanced search strategies of health-related computerized databases. In addition, the search included the bibliographies of key reviews and of all articles that met the search criteria.

Articles were accepted for review if they used an objective gold standard (ie, Manning, Kruis, or Rome I criteria or clinical assessment) and met one of the following criteria: compared 2 diagnostic modalities, distinguished IBS from another condition, appraised individual symptoms for their diagnostic association, or provided sensitivity/specificity data on a diagnostic modality.

**EVALUATION OF PUBLISHED STUDIES**

The symptom criteria were assessed for their test characteristics and performance (Table 1) based on abstraction of the following information from each study: study design (ie, gold standard, diagnostic performance, and disease prevalence) and diagnostic accuracy (ie, sensitivity, specificity, and diagnostic odds ratio).

The quality of each study was assessed by summing the weights of the study characteristics met. These weights were obtained from a multivariate regression analysis reported in a recent publication by Lijmer and colleagues that evaluated design-related bias in assessments of diagnostic tests. The potential range of each study’s total score was from...
RESULTS

The initial search strategy identified 291 titles. One hundred twenty-eight abstracts remained after explicit title rejection criteria were applied. Further evaluation of these 128 abstracts resulted in 88 articles for final review. Of the 88 articles identified, 28 (32%) met the criteria for inclusion in the systematic review.

PERFORMANCE AND QUALITY EVALUATION OF DIAGNOSTIC CRITERIA

Manning Criteria

Of the 3 sets of diagnostic criteria identified in this review, the Manning criteria appeared to be the most extensively studied. Manning and colleagues' originally identified only 4 symptoms that were significantly more prevalent in IBS patients than in organic controls. Two further symptoms approached statistical significance (mucus per rectum and sensation of incomplete evacuation). Using the 4 significant symptoms, subjects having less than 2 symptoms had a positive predictive value for IBS of 12%. If 2 or more symptoms were present, the positive predictive value was 74%. Finally, if 2 or more symptoms were present with all 6 symptoms included, the positive predictive value was 63%. Groups have continued to determine sensitivity and specificity data for 2 of 4 and 2 or 3 of 6 symptoms (Table 3).

When 2 of 4 symptoms were used, the Manning criteria yielded a sensitivity and specificity of 91% and 70%, respectively. In addition, when 2 or more of 6 criteria were used, sensitivity ranged from 84% to 94% and specificity ranged from 55% to 76%. In articles in which 3 or more symptoms were assessed (irrespective of whether it was out of 4 or 6), sensitivity ranged from 63% to 90% and specificity ranged from 70% to 93%.

The diagnostic ability of the Manning criteria also depended on the control group used. All but 1 study compared the ability to distinguish IBS from organic gastrointestinal (GI) disease. However, in...
terpretation of the validation studies is problematic since many of the control group patients experienced or had upper GI symptoms rather than the organic lower GI disorders from which IBS more generally needs to be distinguished. The Manning criteria fared better when used to distinguish patients with IBS from healthy controls (sensitivity, 58%-94%; specificity, 65%-66%; specificity, 86%-93%) than when used to distinguish IBS from organic GI disease (sensitivity, 58%-94%; specificity, 55%-93%).

### Kruis Score

Kruis and colleagues used a point system whereby functional symptoms received positive values and “red flag” symptoms received negative values (Table 3). Based on this point system, using a score of 4 or greater, to identify IBS (or ≥3 symptoms from the list in Table 1), the sensitivity was reported as 64% and the specificity as 99%.10

### Rome I Criteria

Rome I criteria were developed through expert consensus as the first of an ongoing series of criteria for the standardization of diagnostic criteria for IBS.3 The 3 elements of the Manning criteria that were elucidated in factor analysis constitute the first part of the Rome I criteria. Despite this, the published validation of these criteria is minimal. Table 3 summarizes data from 2 studies, only one of which provided an evaluation of the sensitivity and specificity of the criteria.23 The Rome I criteria demonstrated a sensitivity of 65% and specificity of 100%. The positive predictive value ranged between 69% and 100% in these 2 patient groups. However, the study had a relatively small sample size and combined the absence of red flag features with symptom criteria.

### Comparisons of Criteria

Although the Rome I criteria have not been well tested in a controlled fashion, studies have tried to compare results between the various criteria. Two articles compared the agreement among various diagnostic criteria for IBS. There was good agreement between the Manning and Rome I diagnosis of IBS in 1 study (κ = 0.72).24 Additionally, in a large population-based study, 98% of subjects who tested positive for the Rome I criteria also met the Manning criteria. However, of the subjects who tested positive for 2 or more of the Manning criteria, only 37% were positive.

---

**Table 3. Summary of Studies Validating Standard Diagnostic Criteria**

<table>
<thead>
<tr>
<th>Criteria Source, y</th>
<th>Scoring Method</th>
<th>Gold Standard</th>
<th>Type of Control</th>
<th>IBS Patients, No.</th>
<th>Type of Control</th>
<th>Sensitivity, %</th>
<th>Specificity, %</th>
<th>Diagnostic Odds Ratio</th>
<th>Score†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Manning et al,1978</td>
<td>≥2 of 4 criteria</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>33</td>
<td>32</td>
<td>91</td>
<td>70</td>
<td>22.2</td>
<td>6 10.9</td>
</tr>
<tr>
<td>Talley et al,1989</td>
<td>≥2 of 6 criteria</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>33</td>
<td>32</td>
<td>94</td>
<td>55</td>
<td>12.5</td>
<td>6 7</td>
</tr>
<tr>
<td>Talley et al,1991; Talley,1992</td>
<td>≥2 of 4 criteria, ≥2 of 6 criteria</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>154</td>
<td>82</td>
<td>65</td>
<td>86</td>
<td>11.4</td>
<td>6 7</td>
</tr>
<tr>
<td>Jeong et al,1993</td>
<td>≥3 of 3 criteria</td>
<td>Clinical</td>
<td>Non-IBS referrals</td>
<td>114</td>
<td>58</td>
<td>84</td>
<td>54</td>
<td>6.9</td>
<td>7 12.3</td>
</tr>
<tr>
<td>Rao et al,1993</td>
<td>≥3 of 6 criteria</td>
<td>Clinical</td>
<td>Healthy</td>
<td>45</td>
<td>65</td>
<td>66</td>
<td>93</td>
<td>27.4</td>
<td>8 13.2</td>
</tr>
<tr>
<td>Dogan and Unal,1996</td>
<td>≥3 of 6 criteria</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>182</td>
<td>165</td>
<td>90</td>
<td>86</td>
<td>57.4</td>
<td>7 12.3</td>
</tr>
<tr>
<td>Thompson,1997</td>
<td>≥2 of 6 criteria</td>
<td>Clinical</td>
<td>None</td>
<td>201</td>
<td>52</td>
<td>56</td>
<td>95</td>
<td>21.8</td>
<td>8 13.2</td>
</tr>
<tr>
<td>Frigerio et al,1992</td>
<td>Kruis (male) ≥44</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>108</td>
<td>57</td>
<td>67</td>
<td>90</td>
<td>18.7</td>
<td>8 13.2</td>
</tr>
<tr>
<td>Dogan and Unal,1996</td>
<td>Kruis ≥44</td>
<td>Clinical</td>
<td>Organic GI</td>
<td>182</td>
<td>165</td>
<td>81</td>
<td>91</td>
<td>44.8</td>
<td>7 12.3</td>
</tr>
</tbody>
</table>

*IBS indicates irritable bowel syndrome; GI, gastrointestinal; NUD-OGD, nonulcer dyspepsia-oesophagogastroduodenoscopy; and NC, not calculable.
†Sums of scores from Table 1.

©2001 American Medical Association. All rights reserved.

Downloaded From: http://archinte.jamanetwork.com/pdfaccess.ashx?url=/data/journals/intemed/11991/ on 03/31/2017
using the Rome I criteria. The lower prevalence rate could be due to the inclusion of pain as a necessary pre-condition in the Rome I criteria.

Quality of the Criteria

Each of the studies used to validate standard diagnostic criteria was scored based on quality criteria. The raw scores ranged from 1 through 8, and weighted scores ranged from 3 to 13.2. Of the 7 validation studies for the Manning criteria, all but 1 were of medium to high quality. The 3 validation studies on the Kruijssen criteria all received a medium to high quality score. Only 1 of the 2 validation studies for the Rome I criteria obtained a medium to high quality score. Two additional studies were identified that compared diagnostic criteria. However, these studies did not compare the criteria with a diagnostic gold standard.

DIAGNOSTIC ALGORITHM

The diagnostic algorithm was developed based on consensus of the guideline statements that were derived from the systematic review, supplemental review, and expert opinion. The algorithm consisted of a primary module, a primary care workup module that comprised 3 predominant symptom patterns (constipation, diarrhea, and pain), and a subspecialist referral module (Figures 1, 2, and 3).

While the evidence suggests that the Manning criteria have the greatest number of validation studies, the expert panel reached consensus and selected the Rome II criteria as the primary diagnostic symptom criteria. The Rome II criteria incorporate the most valid elements of the Manning criteria while broadening inclusion with the addition of abdominal discomfort or pain and potentially greater discrimination between IBS and other functional disorders. In addition, the subcategorization of IBS on the validated Rome I platform facilitates management in a clinical algorithm. Given the validity of the Manning criteria, the panel alternatively accepted chronic abdominal pain plus 2 or more Manning criteria as an acceptable criterion for the algorithm. Severity of illness was classified into 3 categories: mild (can be ignored if the patient does not think about it), moderate (cannot be ignored but does not affect patient’s lifestyle), and severe/very severe (affects patient’s lifestyle). Predominant symptom patterns were chosen by the expert panel based on the categorization of the Rome Working Group.

After patients are categorized, a thorough history should be taken to identify previous interventions, therapies, and medications used. In some cases, a psychosocial assessment is recommended. The expert panel achieved consensus that an empirical trial of therapy based on the predominant symptom complex does aid in the treatment of patients with suspected IBS. Failure to respond to empirical trials may have diagnostic implications as in other functional GI disorders, although evidence is forthcoming. Conservative empirical trials may include the use of antidiarrheal agents for predominant diarrhea symptoms or antispasmodics for predominant pain symptoms. Trials could also include psychosocial counseling, stress reduction, or biofeedback based on needs assessment. Further diagnostic consideration may be undertaken depending upon responses to this trial. Upon referral to a subspecialist, traditional invasive and noninvasive testing is recommended, if necessary, to establish the diagnosis and to arrive at a therapeutic approach targeted at the predominant symptom complex.

In the development of the algorithm, there were areas of agreement and disagreement that impacted the specific diagnostic
The objective of the present study was to use the best available evidence, supplemented by expert opinion, to arrive at evidence- and consensus-based guidelines for a diagnostic approach to patients with suspected IBS. Scant data were identified in the published literature regarding the effectiveness of competing diagnostic approaches, the accuracy of diagnostic tests, and the internal validity of current diagnostic symptom criteria. As a result, it was necessary to rely upon previously published validation studies, previously developed practice guidelines, and the consensus opinion of our expert panel when developing guidelines. To achieve the study’s objective, a panel was assembled that reflects the needs and concerns of primary care providers.

Preliminary efforts toward the development of diagnostic guidelines entailed the systematic review of previous investigations. Of the studies identified in the review, 8 assessed the Manning criteria, 7,15–21 3 evaluated the Kruis criteria, 10,20,22 and 2 evaluated the Rome I criteria 21,23 (Table 3). The latest Rome II criteria are based on needed improvements to the Rome I criteria and use the most valid elements of the Manning criteria. However, it is evident that more research is needed to better validate both the Rome I and Rome II criteria. Still, the potential advantages of the Rome II criteria include simplicity and improved sensitivity as a result of the inclusion of discomfort and pain as symptoms. Furthermore, the Rome II criteria have potentially greater specificity given that they do not include the second part of the Rome I criteria (non–pain-related symptoms), which had poor clustering in factor analysis.

Over the past decade, research has begun to reveal differences in physiological findings in IBS subjects with different predominant symptom patterns, including dysmotility, gut hypersensitivity, and altered brain activation, among others.11 To subclassify by predominant symptom, diarrhea vs constipation, new criteria were needed to better identify these subgroups. Thus, the Rome II criteria may offer improved discriminative ability for diagnosing patients with IBS.8 While abstracts assessing the validity of the Rome II criteria have been presented, 27–31 no full-length reports on the validity of the Rome II criteria had been published at the time of this report.

The Manning and Rome criteria have gained much attention. However, the Kruis score10 has not been as widely adopted, possibly because of the inclusion of red flag symptoms as part of the scoring algorithm. Red flag symptoms are quite common in subjects with IBS, and, based on the present review, blood in the stool may be seen in up to 31% of IBS subjects, with no objective cause identified on subsequent evaluation.13 Blood in the stool

---

Figure 2. Irritable bowel syndrome diagnostic algorithm for constipation, diarrhea, and abdominal pain (module 1).
alone could represent hemorrhoidal bleeding in IBS patients, yet, in the Kruis score, this would incur a penalty of 98 points—enough to fail to meet the criteria for IBS. This may explain the relatively low sensitivity and high specificity of the Kruis score. Frigerio and colleagues adjusted the score to exclude a diagnosis of organic digestive disease in patients with 44 or more points. Still, their modification was unable to significantly improve the sensitivity of the criteria.

Based on the present evaluation, studies that assessed standard diagnostic criteria were generally of medium to high quality. However, only 1 study of medium to high quality evaluating the Rome I criteria was identified in the review. Although surveys and symptom criteria have been used as an aid to identify IBS and to distinguish IBS from other functional disorders, procedures have most often been relied upon to rule out organic disease. While the expert panel reached agreement regarding the use of sigmoidoscopy and colonoscopy in the guidelines, no validated evidence was found to support the diagnostic value of these and other commonly used invasive and noninvasive procedures (eg, blood tests or colonic transit studies). Moreover, recent evidence suggests that testing for and treating small intestine bacterial overgrowth in IBS may result in improved outcomes, but diagnostic utility in the primary care practice setting requires further validation.

The diagnostic algorithm presented represents an accumulation of the best available evidence- and consensus-based expert opinion from a variety of practice settings. We recognize that there is a lack of expert consensus in many areas—both among experts and between experts and the published literature. Guideline statements were incorporated that were deemed appropriate even if there was disagreement among experts. This enhances the flexibility of the algorithm, allowing providers greater opportunity to employ their own judgment. Accepted components of the algorithm include the differential treatment of patients based on predominant symptom type (constipation vs diarrhea vs abdominal pain) and the fact that younger patients without alarm symptoms should be seen initially in the primary care setting. There is agreement that symptom severity should play a role in the intensity of treatment. However, it is well known that IBS patients often present with extraintestinal symptoms, especially psychological comorbidity, that may dramatically influence the classification of severity. While we advocate empirical trials in our algorithm as an aid to management, the diagnostic validity of this approach remains unclear. Furthermore, the utility of newly available medications targeting the pathophysiological mechanisms of IBS remains unclear. Still, these medications hold promise for more targeted empirical trials based on the pathophysiological mechanism of the predominant symptom complex. The potential utility of targeted empirical trials is that a treatment response may become a diagnostic indicator in itself. Therefore, the predictive value of empirical therapy must be assessed in prospective trials.

There are several limitations to the present study. First, the findings of our systematic review were likely confounded by publication bias; that is, small studies with positive findings are selectively published. Thus, it is possible that studies with negative findings regarding the discriminative capability of symptom criteria or poor test characteristics of standard diagnostic tests may not have been discovered in our review. There were also many gaps in the literature, which meant that expert opinion was required to develop the algorithm. Additionally, several areas of disagreement remained even after a
modified Delphi method was used. The algorithm was based primarily on expert consensus, yet, in some cases, consensus-based recommendations were not possible, as clearly elucidated in the algorithm. As with all guidelines, providers must use their best judgment in determining which patients are eligible for the guidelines and in which cases the guidelines should be strictly adhered to. Finally, because of the scope of the present study, there are no recommendations in the algorithm regarding the possible impact that sex has on IBS symptom reporting and on symptom-based diagnostic criteria. Indeed, differences between the sexes in health-seeking behavior have been reported by Hochstrasser and Angst, who found that women sought care for GI problems significantly more often than did men.

Further research is necessary to define the most accurate methods of patient identification and diagnosis. Studies should be performed using established methodological standards for diagnostic test evaluation and should compare the most commonly used criteria and diagnostic tests. Finally, a prospective evaluation of the impact of systematic approaches to care for patients with IBS should be performed to document the impact of guidelines on the cost-effectiveness and outcomes of care. We hope that, until results from comparative prospective studies are available, the algorithm will inform the decision-making process for a wide range of providers caring for primary care patients with abdominal discomfort or pain and altered bowel function suggestive of IBS.

Accepted for publication March 29, 2001.

From the Department of Medicine, Southern Arizona Veterans Affairs Medical Center, Tucson (Dr Fass); Department of Medicine, Kaiser Permanente Medical Care Program, San Diego, Calif (Drs Longstreth, Crane, and McCarrberg); Departments of Medicine (Drs Pimentel and Ofman) and Health Services Research (Dr Ofman), Cedars-Sinai Medical Center, Los Angeles, Calif; Zynx Health Inc, Los Angeles (Messrs Fullerton and Russak, Drs Chiou and Ofman, and Ms Reyes); and Department of Medicine, Vanderbilt University Medical Center, Nashville, Tenn (Dr Eisen).

This study was sponsored by an educational grant from the Novartis Pharmaceuticals Corp, East Hanover, N J.

Corresponding author and reprints: Joshua Ofman, MD, MSHS, Zynx Health Inc, 9100 Wilshire Blvd, East Tower, Suite 655, Beverly Hills, CA 90212 (e-mail: ofmanj@zynx.com).

REFERENCES


Downloaded From: http://archinte.jamanetwork.com/pdfaccess.ashx?url=/data/journals/intemed/11991/ on 03/31/2017