

# 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; Chiun-Fang Chiou, 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 diag-

nostic 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

**I**RRITABLE 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.<sup>1</sup> The incidence of IBS is reported to be 15% to 20% in the general population,<sup>2,3</sup> with prevalence rates dependent upon the symptom criteria<sup>4</sup> used to define the condition. Furthermore, functional bowel complaints such as IBS are responsible for nearly 50% of visits to gastroenterologists.<sup>5</sup>

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 in-

clude the Rome I criteria,<sup>6</sup> the Manning criteria,<sup>7</sup> and the recently developed Rome II consensus criteria.<sup>8</sup> 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,<sup>9</sup> 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 affiliations of the authors appear in the acknowledgment section at the end of the article.*

**Table 1. Comparison of Diagnostic Criteria\***

| Manning Criteria†  | Kruis Questionnaire‡  | Rome I Criteria§   | Rome II Criteria  |
|--|---|--|---|
| Abdominal distension<br>Pain relief with bowel movement<br>More frequent stools with pain<br>Looser stools with pain<br>Mucus per rectum¶<br>Feeling of incomplete evacuation¶ | <p><b>Patient fills out:</b></p> <ol style="list-style-type: none"> <li>Did you come because of abdominal pain?<br/>Do you suffer from flatulence?<br/>Do you suffer from irregularities of bowel movements?</li> <li>Have you suffered from your complaints for more than 2 years?</li> <li>How can your abdominal pain be described: burning, cutting, very strong, terrible, feeling of pressure, dull, boring, not so bad?</li> <li>Have you ever noticed alternating constipation and diarrhea?</li> </ol> <p><b>Physician fills out:</b></p> <ol style="list-style-type: none"> <li>Abnormal physical findings and/or history pathognomonic for any diagnosis other than IBS</li> <li>ESR &gt;20 mm/2 h</li> <li>Leukocytosis (&gt;10 000/cm<sup>3</sup>)</li> <li>Hemoglobin: female, &lt;12 g/dL; male, &lt;14 g/dL</li> <li>History of blood in stool</li> </ol> | <p><b>&gt;3 mo of continuous or recurrent symptoms:</b></p> <p>Abdominal pain:<br/>Relieved with defecation<br/>and/or<br/>Associated with change in stool consistency</p> <p><b>Plus</b></p> <p>≥2 of the following on at least 25% of days:<br/>Altered stool frequency (&gt;3/d or &lt;3/wk)<br/>Altered stool form (lumpy/hard or loose/watery)<br/>Altered stool passage (straining, urgency, or feeling of incomplete evacuation)<br/>Passage of mucus<br/>Bloating or feeling of abdominal distension</p> | <p><b>≥12 wk of continuous or recurrent symptoms:</b></p> <p>Abdominal pain or discomfort</p> <p><b>Plus</b></p> <p>≥2 of the following:<br/>Relieved with defecation<br/>and/or<br/>Associated with change in frequency of stool<br/>and/or<br/>Associated with change in form (appearance) of stool</p> |

\*IBS indicates irritable bowel syndrome; ESR, erythrocyte sedimentation rate.

†From Manning et al.<sup>7</sup>

‡From Kruis et al.<sup>10</sup> Questions that are not associated with a score are excluded.

§From Drossman et al.<sup>6</sup>

||From Thompson et al.<sup>11</sup>

¶Differences between IBS patients and controls were not statistically significant ( $P > .05$ ).

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 accu-

racy 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 assess-

ment) 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<sup>1</sup> (ie, gold standard, diagnostic performance, and disease prevalence) and diagnostic accuracy<sup>2</sup> (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<sup>12</sup> that evaluated design-related bias in assessments of diagnostic tests. The potential range of each study's total score was from

0 to 8 unweighted and from 0 to 13.2 weighted. The study characteristics included spectrum (clinical population vs case-control), verification (complete vs different reference tests vs partial), interpretation of test results (blinded vs not blinded), patient selection (consecutive vs nonconsecutive), data collection (prospective vs retrospective vs unknown), details test (sufficient vs insufficient), details reference test (sufficient vs insufficient), and details population (sufficient vs insufficient). Studies were classified as low quality if they scored in the lowest tertile. Conversely, studies were classified as being of medium-high quality if their scores were in the middle or high tertile (**Table 2**).

### EXPERT PANEL

Because many areas were not addressed in the published literature, in order to complete a diagnostic algorithm, an expert panel was assembled consisting of physicians from 3 medical settings: an academic medical center, a Veterans Affairs medical center, and a large group-model health maintenance organization. Guideline statements were developed from the systematic review, existing guidelines identified in the supplemental literature review, and expert opinion. The panel was asked to vote on the relative appropriateness of each guideline statement by considering both the expected costs and health benefits. Response options were based on a scale from 1 to 9, ranging from extremely inappropriate<sup>1</sup> to extremely appropriate.<sup>9</sup> A score of 5 was considered equivocal. The RAND appropriateness methodology for scoring responses<sup>13</sup> was used as the basis for determining consensus. Results were evaluated with respect to tertile<sup>1-9</sup> after discarding the lowest and highest scores. Agreement was reached if the remaining 4 scores fell within any 3-point range. Disagreement occurred if at least 1 of the remaining 4 ratings fell within the lowest tertile and at least 1 score fell within the highest tertile. Unclear opinion was defined as all of the votes falling within adjoining tertiles. Experts were allowed to modify their votes after independently reviewing the results of the group's ratings.

### DIAGNOSTIC ALGORITHM

Guidelines in the form of a diagnostic algorithm were developed incorporating the best available evidence and expert opinion regarding the diagnosis of IBS. Expert opinion was used when evidence in the literature did not exist to inform the decision. Guideline statements were incorporated into the algorithm when they met 1 of 2 criteria: if there was strong literature-based evidence or if the expert panel voted that the guideline statement was appropriate (ie, a median score >5). The algorithm details whether there was agreement or disagreement among the expert panel. The methodology of the algorithm flowcharts is consistent with the format previously adopted by the Agency for Health Care Policy Research.<sup>14</sup>

## RESULTS

### SYSTEMATIC REVIEW

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<sup>7</sup> 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

**Table 2. Criteria for Grading the Quality of Diagnostic Studies**

| Study Characteristic           | Score |           |
|--------------------------------|-------|-----------|
|                                | Raw   | Weighted* |
| Spectrum                       |       |           |
| Clinical population            | 1     | 3.0       |
| Case-control                   | 0     | 0         |
| Verification                   |       |           |
| Complete                       | 1     | 3.2       |
| Different reference tests      | 0     | 1.0       |
| Partial                        | 0     | 0         |
| Interpretation of test results |       |           |
| Blinded                        | 1     | 1.3       |
| Not blinded                    | 0     | 0         |
| Patient selection              |       |           |
| Consecutive                    | 1     | 0.9       |
| Nonconsecutive                 | 0     | 0         |
| Data collection                |       |           |
| Prospective                    | 1     | 1.0       |
| Unknown                        | 0     | 0.5       |
| Retrospective                  | 0     | 0         |
| Details test                   |       |           |
| Sufficient                     | 1     | 1.7       |
| Insufficient                   | 0     | 0         |
| Details reference test         |       |           |
| Sufficient                     | 1     | 0.7       |
| Insufficient                   | 0     | 0         |
| Details population             |       |           |
| Sufficient                     | 1     | 1.4       |
| Insufficient                   | 0     | 0         |

\*The values used to develop the weights for the weighted score are based on results reported by Lijmer et al.<sup>12</sup> Data that were not reported in the article by Lijmer et al are scored as 0.

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.<sup>7,16,17</sup> In addition, when 2 or more of 6 criteria were used, sensitivity ranged from 84% to 94% and specificity ranged from 55% to 76%.<sup>7,16-18,21</sup> 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%.<sup>16-20</sup>

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.<sup>15</sup> However, in-

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

| Criteria                           | Source, y  | Scoring Method     | Gold Standard | Type of Control   | Controls, No. | IBS Patients, No. | Sensitivity, % | Specificity, % | Diagnostic Odds Ratio | Score† |          |    |      |
|------------------------------------|--|--------------------|---------------|-------------------|---------------|-------------------|----------------|----------------|-----------------------|--------|----------|----|------|
|                                    |  |                    |               |                   |               |                   |                |                |                       | Raw    | Weighted |    |      |
| Manning                            | Manning et al, <sup>7</sup> 1978                             | ≥2 of 4 criteria   | Clinical      | Organic GI        | 33            | 32                | 91             | 70             | 22.2                  | 6      | 10.9     |    |      |
|                                    |  |                    | Clinical      | Organic GI        | 33            | 32                | 94             | 55             | 12.5                  |        |          |    |      |
|                                    | Talley et al, <sup>15</sup> 1989                             | No. of warnings    | Clinical      | Healthy           | 154           | 82                | 65             | 86             | 11.4                  |        |          |    |      |
|                                    |  |                    |               | Organic GI        | 101           | 82                | 58             | 74             | 3.9                   |        |          |    |      |
|                                    |  |                    |               | NUD-OGD           | 134           | 82                | 42             | 85             | 4.1                   |        |          |    |      |
|                                    | Talley et al, <sup>16</sup> 1991; Talley, <sup>17</sup> 1992 | ≥2 of 4 criteria   | Clinical      | Organic GI        | 33            | 32                | 91             | 70             | NC                    |        |          |    |      |
|                                    |  |                    |               |                   |               |                   | 94             | 55             | NC                    |        |          |    |      |
|                                    |  | ≥2 of 6 criteria   |               |                   |               |                   | 63             | 85             | NC                    |        |          |    |      |
|                                    |  |                    |               |                   |               |                   | 84             | 76             | NC                    |        |          |    |      |
|                                    | Jeong et al, <sup>18</sup> 1993                              | ≥3 of 4 criteria   | Clinical      | Non-IBS referrals | 114           | 58                | 84             | 54             | 6.5                   |        |          |    |      |
| 67                                 |  |                    |               |                   |               |                   | 70             | 4.8            |                       |        |          |    |      |
| Rao et al, <sup>19</sup> 1993      | ≥3 of 6 criteria   | Clinical           | Healthy       | 45                | 65            | 66                | 93             | 27.4           |                       |        |          |    |      |
|                                    |  |                    |               |                   |               | 23                | 65             | 66             | 61                    | 3.0    |          |    |      |
|                                    |  |                    |               |                   |               | 35                | 65             | 66             | 91                    | 20.9   |          |    |      |
| Dogan and Unal, <sup>20</sup> 1996 | ≥3 of 6 criteria   | Clinical           | Organic GI    | 182               | 165           | 90                | 86             | 57.4           | 7                     | 12.3   |          |    |      |
| Thompson, <sup>21</sup> 1997       | ≥2 of 6 criteria   | Clinical           | None          | NC                | 156           | NC                | NC             | NC             | 1                     | 3      |          |    |      |
| Kruis                              | Kruis et al, <sup>10</sup> 1984                              | Kruis ≥44          | Clinical      | Organic GI        | 209           | 108               | 64             | 99             | 183.1                 | 8      | 13.2     |    |      |
|                                    |  |                    |               |                   |               |                   |                |                |                       |        |          |    |      |
|                                    | Frigerio et al, <sup>22</sup> 1992                           | Kruis (male) ≥44   | Clinical      | Organic GI        | 93            | 15                | 47             | 94             | 12.7                  |        |          |    |      |
|                                    |  |                    |               |                   |               |                   | 108            | 37             | 59                    |        |          | 95 | 30.2 |
|                                    |  |                    |               |                   |               |                   | 201            | 52             | 56                    |        |          | 95 | 21.8 |
|                                    |  | Kruis (female) ≥44 |               |                   |               |                   | 67             | 90             | 18.7                  |        |          |    |      |
| 68                                 |  |                    |               |                   |               |                   | 92             | 23.8           |                       |        |          |    |      |
|                                    | Modified Kruis (male) ≥44                                    |                    |               | 201               | 52            | 68                | 91             | 21.5           |                       |        |          |    |      |
| Dogan and Unal, <sup>20</sup> 1996 | Kruis ≥44  | Clinical           | Organic GI    | 182               | 165           | 81                | 91             | 44.8           | 7                     | 12.3   |          |    |      |
| Rome                               | Thompson, <sup>21</sup> 1997                                 | Rome guidelines    | Clinical      | None              | NC            | 156               | NC             | NC             | NC                    | 1      | 3        |    |      |
|                                    | Vanner et al, <sup>23</sup> 1999                             | Rome guidelines    | Clinical      | Non-IBS referrals | 52            | 46                | 65             | 100            | 18.8                  | 6      | 9        |    |      |

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

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, 65%-66%; specificity, 86%-93%)<sup>15,19</sup> than when used to distinguish IBS from organic GI disease (sensitivity, 58%-94%; specificity, 55%-93%).<sup>7,15-17,19,20</sup>

### Kruis Score

Kruis and colleagues<sup>10</sup> 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 44 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%.<sup>10</sup>

### 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.<sup>5,8</sup> 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.<sup>23</sup> The Rome I criteria demonstrated a sensitivity of 65% and specificity of 100%. The positive predictive value ranged be-

tween 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 ( $\kappa=0.72$ ).<sup>24</sup> Additionally, in a large population-based study, 98% of subjects who tested positive for the Rome I criteria also met the Manning criteria.<sup>25</sup> However, of the subjects who tested positive for 2 or more of the Manning criteria, only 37% were positive

using the Rome I criteria. The lower prevalence rate could be due to the inclusion of pain as a necessary precondition in the Rome I criteria.

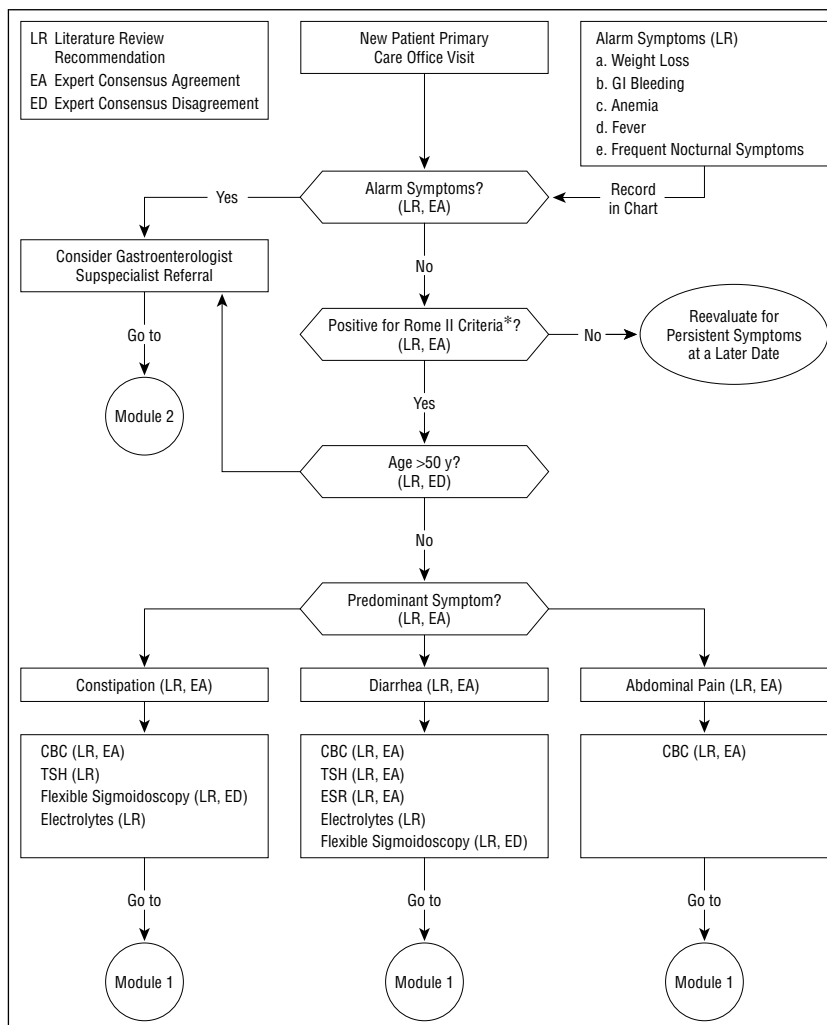
### Quality of the Criteria

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 Kruis 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.<sup>24,25</sup> 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



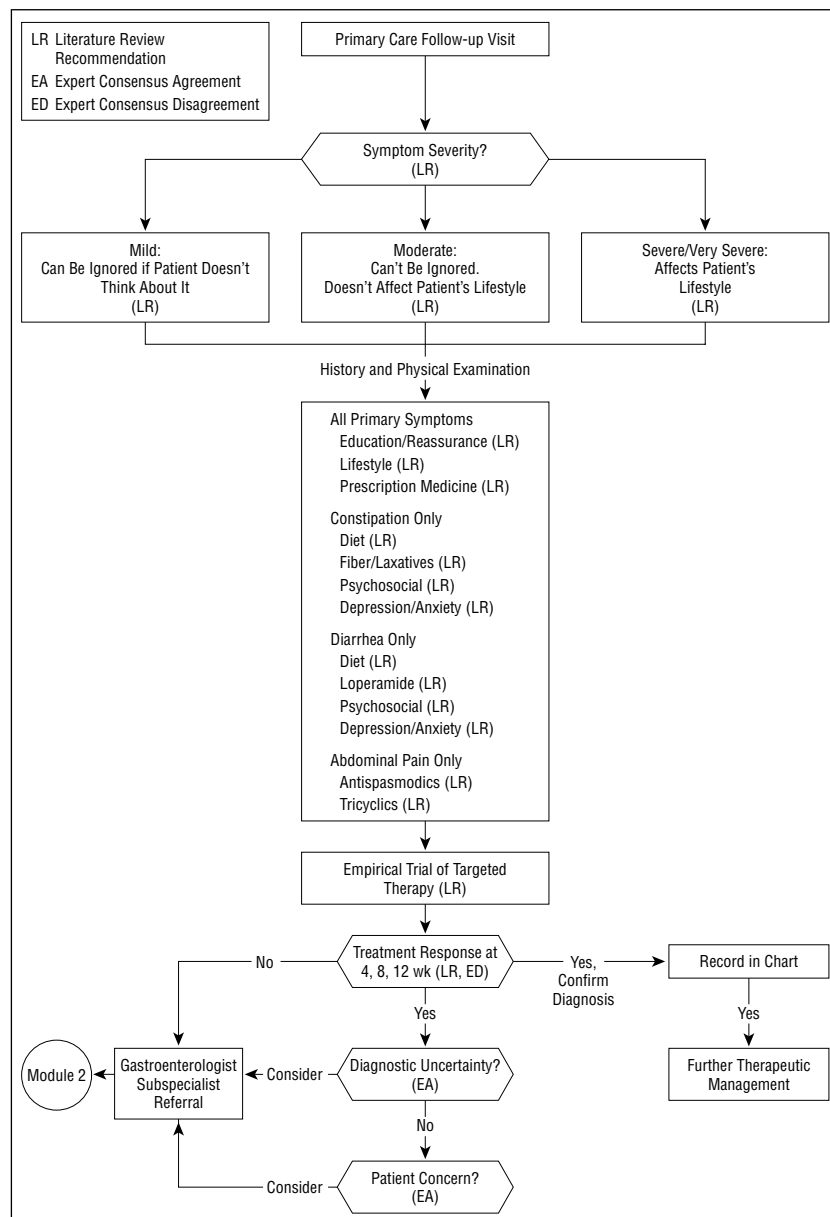
**Figure 1.** Irritable bowel syndrome diagnostic algorithm for a new primary care patient. Asterisk indicates that experts agree that chronic abdominal pain plus 2 or more Manning criteria is an acceptable alternative to the Rome I criteria. GI indicates gastrointestinal; CBC, complete blood cell count; TSH, thyrotropin (thyroid-stimulating hormone) level; and ESR, erythrocyte sedimentation rate.

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).<sup>26</sup> Predominant symptom patterns were chosen by the expert panel based on the categorization of the Rome Working Group.<sup>11</sup>

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



**Figure 2.** Irritable bowel syndrome diagnostic algorithm for constipation, diarrhea, and abdominal pain (module 1).

approach. For example, all of the experts agreed that it is “inappropriate” to have all patients who are “. . . referred to the gastroenterology subspecialty unit with suspected IBS be given an anorectal manometry exam.” Conversely, experts disagreed whether it would be appropriate that all patients “. . . referred to the gastroenterology subspecialty unit with suspected IBS be given a large bowel exam.”

#### COMMENT

The objective of the present study was to use the best available evi-

dence, 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 ob-

jective, 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,<sup>7,15-21</sup> 3 evaluated the Krus criteria,<sup>10,20,22</sup> and 2 evaluated the Rome I criteria<sup>21,23</sup> (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.

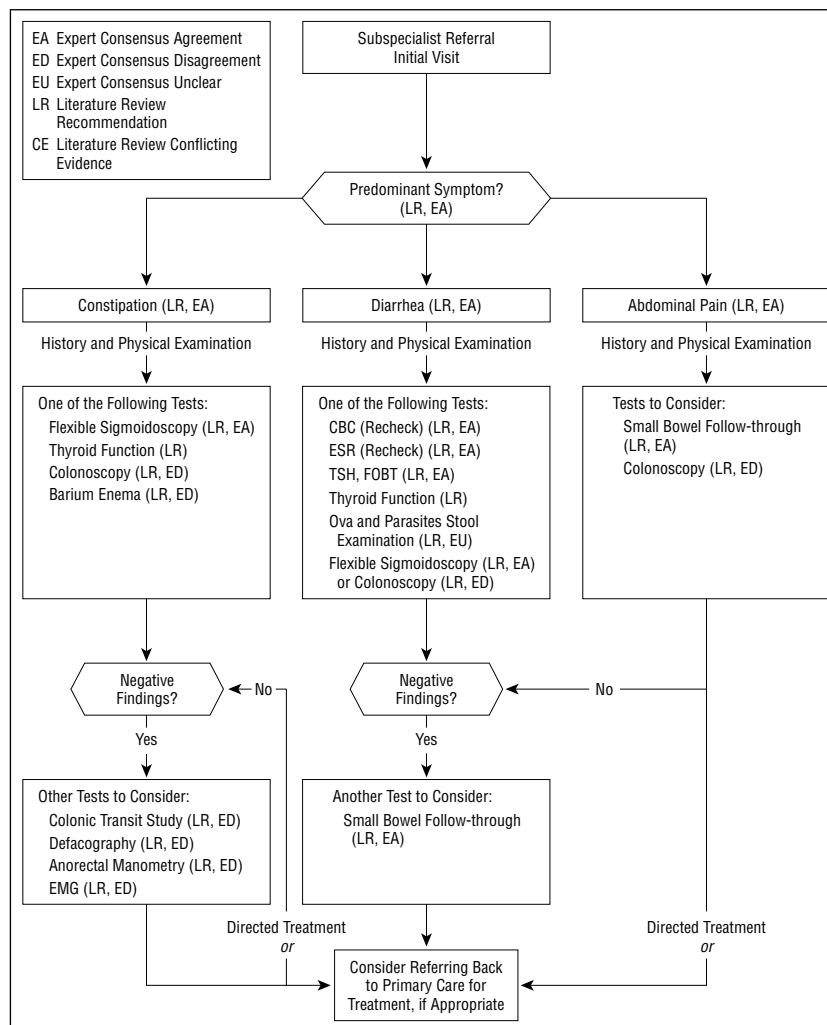
Over the past decade, research has begun to reveal differences in physiological findings in IBS subjects with different predominant symptom patterns, including dysmotility, gut hypersensitivity, and altered brain activation, among others.<sup>11</sup> To subclassify by predominant symptom, diarrhea vs 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.<sup>8</sup> While abstracts assessing the validity of the Rome II criteria have been presented,<sup>27-31</sup> 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 Krus score<sup>10</sup> 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.<sup>15</sup> Blood in the stool

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<sup>22</sup> 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,<sup>26,32-35</sup> 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,<sup>36</sup> 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)



**Figure 3.** Irritable bowel syndrome diagnostic algorithm for gastrointestinal subspecialty workup (module 2). CBC indicates complete blood cell count; ESR, erythrocyte sedimentation rate; TSH, thyrotropin (thyroid-stimulating hormone) level; FOBT, fecal occult blood test; and EMG, electromyography.

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 util-

ity 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,<sup>37</sup> 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 McCarberg); Departments of Medicine (Drs Pimentel and Ofman) and Health Services Research (Dr Ofman), Cedars-Sinai Health System, Los Angeles, Calif; Zynx Health Inc, Los Angeles (Messrs Fullerton and Russak, Drs Chiou and Ofman, and Ms Reyes); and Depart-

ment 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, NJ.

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

- Camilleri M, Choi MG. Review article: irritable bowel syndrome [review]. *Aliment Pharmacol Ther.* 1997;11:3-15.
- Drossman DA, Whitehead WE, Camilleri M. Irritable bowel syndrome: a technical review for practice guideline development [review]. *Gastroenterology.* 1997;112:2120-2137.
- Thompson WG, Heaton KW. Functional bowel disorders in apparently healthy people. *Gastroenterology.* 1980;79:283-288.
- Saito YA, Locke GR, Talley NJ, Zinsmeister AR, Fett SL, Melton LJ III. 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.
- Thompson WG. The functional gastrointestinal bowel disorders. In: Drossman D, ed. *The Functional Gastrointestinal Disorders*. McLean, Va: Degnon Associates; 1994:117-134.
- Drossman DA, Richter JE, Talley NJ. *The Functional Gastrointestinal Disorders, Pathophysiology, and Treatment: A Multinational Consensus*. Boston, Mass: Little Brown; 1994.
- Manning AP, Thompson WG, Heaton KW, Morris AF. Towards positive diagnosis of the irritable bowel. *BMJ.* 1978;2:653-654.
- Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Muller-Lissner SA. Functional bowel disorders and functional abdominal pain. *Gut.* 1999;45(suppl II):II43-II47.
- Schmulson MW, Chang L. Diagnostic approach to the patient with irritable bowel syndrome [review]. *Am J Med.* 1999;107(5A):20S-26S.
- Kruis W, Thieme C, Weinzierl M, Schussler P, Holl J, Paulus W. A diagnostic score for the irritable bowel syndrome: its value in the exclusion of organic disease. *Gastroenterology.* 1984;87:1-7.
- Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Muller-Lissner SA. C. Functional bowel disorders and functional abdominal pain. In: Drossman AD, Talley NJ, Corazziari ETNJ, Thompson WG, Whitehead WE, eds. *Rome II: The Functional Gastrointestinal Disorders*. McLean, Va: Degnon Associates; 2000:351-432.
- Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests [published erratum appears in *JAMA.* 2000;283:1963]. *JAMA.* 1999;282:1061-1066.
- Kahn KL, Roth CP, Kosecoff J, et al. *Indications for Selected Medical and Surgical Procedures—A Literature Review and Ratings of Appropriateness*. Santa Monica, Calif: RAND Corp; 1986.
- Using Clinical Practice Guidelines to Evaluate Quality of Care*. Vol 1. Bethesda, Md: US Dept of Health and Human Services; 1995. Publication AHCPR 95-0045.
- Talley NJ, Phillips SF, Melton LJ, Mulvihill C, Wiltgen C, Zinsmeister AR. Diagnostic value of the Manning criteria in irritable bowel syndrome. *Gut.* 1990;31:77-81.
- Talley NJ, Zinsmeister AR, Van Dyke C, Melton LJ III. Epidemiology of colonic symptoms and the irritable bowel syndrome [published erratum ap-
- pears in *Gastroenterology.* 1992;102:746]. *Gastroenterology.* 1991;101:927-934.
- Talley NJ. Sensitivity and specificity of the Manning criteria [letter]. *Gastroenterology.* 1992;102:1828-1829.
- Jeong H, Lee HR, Yoo BC, Park SM. Manning criteria in irritable bowel syndrome: its diagnostic significance. *Korean J Intern Med.* 1993;8:34-39.
- Rao KP, Gupta S, Jain AK, Agrawal AK, Gupta JP. Evaluation of Manning's criteria in the diagnosis of irritable bowel syndrome. *J Assoc Physicians India.* 1993;41:357-363.
- 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.
- Thompson WG. Gender differences in irritable bowel symptoms. *Eur J Gastroenterol Hepatol.* 1997;9:299-302.
- Frigerio G, Beretta A, Orsenigo G, Tadeo G, Imperiali G, Minoli G. Irritable bowel syndrome: still far from a positive diagnosis. *Dig Dis Sci.* 1992;37:164-167.
- Vanner SJ, Depew WT, Paterson WG, et al. Predictive value of the Rome criteria for diagnosing the irritable bowel syndrome. *Am J Gastroenterol.* 1999;94:2912-2917.
- Kay L, Jorgensen T, Lannig C. Irritable bowel syndrome: which definitions are consistent? *J Intern Med.* 1998;244:489-494.
- Hahn BA, Saunders WB, Maier WC. Differences between individuals with self-reported irritable bowel syndrome (IBS) and IBS-like symptoms. *Dig Dis Sci.* 1997;42:2585-2590.
- Talley NJ, Phillips SF, Wiltgen CM, Zinsmeister AR, Melton LJ III. Assessment of functional gastrointestinal disease: the bowel disease questionnaire. *Mayo Clin Proc.* 1990;65:1456-1479.
- Levy RL, Whitehead WE, Feld AD, et al. Agreement between clinician recorded diagnoses and Rome I and Rome II diagnostic criteria for irritable bowel syndrome [abstract]. *Gastroenterology.* 2000;118:A2068.
- Yuri A, Locke G, Prather C, et al. Evaluation of the Rome II criteria for irritable bowel syndrome in a tertiary care outpatient population [abstract]. *Gastroenterology.* 2000;118:A2071.
- Thompson WG, Irvine JE, Pare P. Comparing Rome I and Rome II criteria for irritable bowel syndrome (IBS) in a prospective survey of the Canadian population [abstract]. *Am J Gastroenterol.* 2000;96:A2553.
- Saito YA, Locke GR, Talley NJ, Zinsmeister AR, Fett SL, Melton J. The effect of new diagnostic criteria for irritable bowel syndrome on community prevalence estimates [abstract]. *Gastroenterology.* 2000;118:A396.
- Sperber A, Safieh Y, Jaffer A, et al. A comparison of the prevalence of IBS using Rome I and Rome II criteria in an epidemiology study. [abstract] *Gastroenterology.* 2000;118:A397.
- Kapoor KK, Nigam P, Rastogi CK, Kumar A, Gupta AK. Clinical profile of irritable bowel syndrome. *Indian J Gastroenterol.* 1985;4:15-16.
- Talley NJ, Phillips SF, Bruce B, Twomey CK, Zinsmeister AR, Melton LJ III. Relation among personality and symptoms in nonulcer dyspepsia and the irritable bowel syndrome. *Gastroenterology.* 1990;99:327-333.
- O'Keefe EA, Talley NJ, Tangalos EG, Zinsmeister AR. A bowel symptom questionnaire for the elderly. *J Gerontol.* 1992;47:M116-M121.
- Agreus L, Talley NJ, Svardsudd K, Tibblin G, Jones MP. Identifying dyspepsia and irritable bowel syndrome: the value of pain or discomfort, and bowel habit descriptors. *Scand J Gastroenterol.* 2000;35:142-151.
- Pimentel M, Chow EJ, Lin HC. Eradication of small intestinal bacterial overgrowth reduces symptoms of irritable bowel syndrome. *Am J Gastroenterol.* 2000;95:3503-3506.
- Hochstrasser B, Angst J. The Zurich Study: XXII, epidemiology of gastrointestinal complaints and comorbidity with anxiety and depression. *Eur Arch Psychiatry Clin Neurosci.* 1996;246:261-272.