Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis

Gea A. Holtman, MSc,† Yvonne Lisman-van Leeuwen, PhD,‡ Johannes B. Reitsma, MD, PhD,§ Marjolein Y. Berger, MD, PhD

abstract

BACKGROUND: The clinical presentation of pediatric inflammatory bowel disease (IBD) is often nonspecific and overlaps with functional gastrointestinal disorders.

OBJECTIVE: To determine the diagnostic accuracy of symptoms, signs, noninvasive tests, and test combinations that can assist the clinician with the diagnosis of IBD in symptomatic children.

METHODS: A literature search was conducted of Medline and Embase. Two reviewers independently selected studies reporting on the diagnostic accuracy of tests for IBD, with confirmation by endoscopy and histopathology or clinical follow-up, in children with chronic gastrointestinal symptoms. Two reviewers independently extracted data and assessed study quality with the QUADAS-2, an evidence-based quality assessment tool for diagnostic accuracy studies.

RESULTS: Nineteen studies were included (N = 2806). Symptoms (abdominal pain, diarrhea, rectal bleeding, and weight loss) had pooled sensitivities ranging from 0.48 to 0.82 and specificities ranging from 0.17 to 0.78. Of all the blood markers, C-reactive protein (CRP) (9 studies) and albumin (6 studies) had the best performance, with pooled sensitivities of 0.63 (0.51–0.73) and 0.48 (0.31–0.66), respectively, and specificities of 0.88 (0.80–0.93) and 0.94 (0.86–0.98). Assessment of fecal calprotectin (FCal) (10 studies) had a pooled sensitivity of 0.99 (0.92–1.00) and a specificity of 0.65 (0.54–0.74). One limitation was that none of the studies was conducted in nonreferred children.

CONCLUSIONS: In children whose pediatrician is considering an endoscopy, symptoms are not accurate enough to identify low-risk patients in whom an endoscopy can be avoided. FCal, CRP, and albumin findings are potentially of clinical value, given their ability to select children at low risk (negative FCal test result) or high risk (positive CRP or albumin test result) for IBD.
Chronic gastrointestinal symptoms in children are a common reason to visit a physician. Differentiation between functional bowel disorders, in which diagnostic testing should be minimized, and inflammatory bowel disease (IBD), which should not be missed, is a diagnostic challenge. The incidence of pediatric IBD is low (5.2 per 100,000 per year), although increasing. Symptoms of IBD are often atypical. Only 25% of children with Crohn’s disease present with the classic triad of symptoms: diarrhea, abdominal pain, and weight loss. Testing all children with chronic gastrointestinal symptoms for IBD is neither necessary nor efficient, particularly not in primary care. Furthermore, to confirm or rule out IBD, an endoscopy is necessary, which is invasive and requires general anesthesia when performed in children. In contrast, an attitude of “wait and see” may cause unnecessary concerns and loss of well-being in children with IBD.

Noninvasive tests for IBD, such as blood markers, fecal markers, and ultrasonography, may assist the clinician with this diagnostic dilemma. These tests can be used as a triage instrument; they assist in safely ruling out existing IBD and in selecting those patients who are candidates for further investigations. Fecal calprotectin (FCal), an inflammatory marker, has been extensively studied in several reviews and meta-analyses and has good properties for ruling out IBD in children presenting to the pediatrician with symptoms suggestive of IBD. However, a complete overview of the diagnostic accuracy of all symptoms, signs, and noninvasive tests is lacking. Moreover, the optimal combination of tests or the additional value of a single test to symptoms or other tests has rarely been studied.

The performance of symptoms, signs, and noninvasive tests may vary between nonreferred and referred children due to a diverse patient mix and underlying disorders, as well as different moments in the course of disease at which patients present or varying reference tests on which a diagnosis is based. The goal of the present study was to systematically review the literature to provide an overview of the accuracy of single symptoms, signs, and noninvasive tests for IBD diagnosed via endoscopy in children presenting with chronic gastrointestinal symptoms in all health care settings. A secondary goal was to present the accuracy of test combinations and the added value of a single test to symptoms, signs, or other tests.

METHODS

Search Strategy

A literature search for eligible diagnostic studies was conducted in Medline and Embase (from inception to September 18, 2014) using Medical Subject Headings, Emtree terms, and free text words related to child, the target condition (IBD), and diagnostic accuracy (Supplemental Information). A search strategy was constructed specific for diagnostic accuracy studies based on published search strategies. An information expert assisted the search. In addition, 2 authors (Y.L.-V.L. and G.A.H.) hand-searched the references of all included full-text articles, 3 systematic reviews, and guidelines on pediatric IBD. No language restrictions were applied to the searches.

Study Selection

We identified studies performed in all health care settings. Six criteria were used to choose studies: (1) the study population consisted of children with gastrointestinal symptoms suggestive of IBD (studies including healthy control subjects and/or patients with known IBD were excluded); (2) one of the following diagnostic tests was investigated: signs, symptoms, markers (blood, fecal, or urinary), or ultrasonography; (3) the reference standard for IBD was endoscopy, including histopathology and/or clinical follow-up; (4) the target condition was IBD; (5) the study design provided information about the association between tests of interests and the presence or absence of IBD; and (6) the study report, or the subsequent data requested, enabled the construction of a 2 × 2 table. Authors were contacted if data for the 2 × 2 table were insufficient or missing.

Two reviewers (Y.L.-V.L. and G.A.H.) independently screened titles and abstracts of all identified articles and assessed full-text articles of each potentially eligible study for inclusion. Disagreements between the reviewers was resolved by discussion and, if necessary, by a third reviewer (M.Y.B.). If the full text of an included study was not available, the first or last author was contacted.

Data Extraction and Quality Assessment

Two reviewers (Y.L.-V.L. and G.A.H.) independently performed data extraction and quality assessment by using standardized forms. Disagreements between the reviewers were resolved by consensus or by a third reviewer (M.Y.B.). The following data were extracted: setting and design; study population; index test; reference standard; prevalence of IBD in the study population; number of patients with Crohn’s disease, ulcerative colitis, IBD unclassified, or no IBD; and data for the 2 × 2 table. Study quality was assessed by using an evidence-based quality assessment tool for diagnostic accuracy studies (the QUADAS-2). Scores for low or high risk of bias were allocated to 4 domains: patient selection, index test, reference standard, and flow and timing (Fig 1). In addition, concerns were scored
regarding applicability for the first 3 domains.

**Data Synthesis and Analysis**

Diagnostic 2 × 2 tables were imported in Review Manager 5.0 (RevMan, Cochrane Collaboration), and sensitivity, specificity, and corresponding 95% confidence intervals (CIs) were calculated for each symptom, sign, test, and test combination. The added value of tests was described when it was reported in the studies. For the meta-analysis, bivariate random effects models were used to calculate pooled estimates of sensitivity, specificity, and likelihood ratios when ≥5 studies per index test were included. The MIDAS module was used for meta-analysis of diagnostic test accuracy studies in Stata/SE version 12.1 (Stata Corp, College station, TX).

**Sources of Heterogeneity**

We evaluated whether differences in certain factors could explain identified heterogeneity. These factors included the following: design (cohort or case-control); setting (according to level of selection, 3 settings were defined [children presenting for the first time in primary care (nonreferred, low risk); children referred by their primary care physician (either primary care physician or pediatrician) to a pediatrician or pediatric gastroenterologist for diagnostic evaluation (referred, moderate risk) and children referred by a pediatrician to a pediatric gastroenterologist and endoscopy (referred, high risk)]; number/choice of reference standards (1 or 2, endoscopy or follow-up); prevalence; and cutoff value of the index test. In case of outliers, we evaluated whether bias or specific study characteristics could explain the result. A subgroup analysis (≥5 studies per subgroup) or sensitivity analysis without outliers was performed to evaluate the effect of heterogeneity on test characteristics.

**Potential Clinical Impact**

Our goal was to provide more insight into the potential clinical consequences of using the results of the investigated tests. For each test for which we were able to calculate pooled sensitivity and specificity, hypothetical 2 × 2 tables were constructed in 100 children with gastrointestinal symptoms. The number of children with IBD was based on the mean IBD prevalence in the cohort studies included in the meta-analysis. By standardizing the prevalence, it is possible to compare the results of each test. The 2 × 2 tables were based on the pooled estimates of sensitivity and specificity of the index test. Clinical impact was interpreted as follows: children with IBD missed are those with IBD and a negative index test result; the numbers of unnecessary endoscopies are the children without IBD with a positive index test result; and the reduction of patients requiring endoscopy is the total number of patients with a negative index test result. For calculating
the latter, we assumed that in the alternative strategy, all 100 hypothetical children would undergo endoscopy.

RESULTS

Selection, Characteristics, and Quality of Studies

The literature search yielded 19 diagnostic studies involving a total of 2806 children with gastrointestinal symptoms (age range: 3 months–21 years), of whom 1265 had IBD (Fig 2). The mean prevalence of IBD in the cohort studies was 54% (range: 19%–82%).15–28 The characteristics of the 14 cohort studies15–28 and 5 case-control studies29–33 are presented in Supplemental Tables 3 and 4. We requested data of incorrect19 or insufficient18,20,22,24–28,33 2 × 2 tables for index tests, and additional data were received from 5 studies.22,24,26–28 Three cohort studies used 2 reference standards: endoscopy and follow-up.25,27,33 None of the studies was performed in nonreferred children (low risk). One study included children referred by family physicians, but the high prevalence of IBD (62%) indicated to us that the children were probably not selected consecutively, and we therefore scored the risk as moderate or high.24 In 4 studies, the referred children were at moderate or high risk23,25,27,33; in 13 studies, they were at high risk7,15–22,26,28,31,32; and in 1 study, the setting was unclear.29 Two case-control studies reported solely on Crohn’s disease.31,33 Table 1 presents the risk of bias of all studies. Seventeen studies found a high risk of bias in 1 domain. On average, the reviewers resolved the disagreement on 2 of 14 items per study (range: 0–5).

Diagnostic Accuracy

Diagnostic accuracy measures of all symptoms, signs, tests, and test combinations evaluated are presented in Supplemental Figure 3 A–E. Table 2 presents the results of the meta-analysis of symptoms, signs, and tests evaluated in ≥5 studies. Setting (moderate/high versus high risk), prevalence, and number of reference standards varied little between studies and could not explain the heterogeneity in test characteristics of any of the symptoms, signs, or tests evaluated.

Symptoms and Signs

The sensitivity and specificity varied substantially between studies for all symptoms (8 studies) (Supplemental Figure 3A). Study design could not explain heterogeneity. Rectal bleeding had the highest positive likelihood ratio of 2.6 (1.7–4.0).

Blood Markers

Sensitivity varied considerably within each blood marker studied (in total, 10 blood markers studied in 13 studies) (Supplemental Figure 3B). Specificity was fairly homogeneous within all blood markers. C-reactive protein (CRP) (cutoff range: 3–10 mg/L) was evaluated in 9 studies.16,19,22,24,26,27,29,30,32 Two studies had high sensitivities and high specificities compared with the other studies, in which only specificity was high.16,29 Heterogeneity could not be explained by differences in study design or cutoff value, nor could we identify specific reasons for bias. Pooled sensitivities for CRP with and without these 2 outliers were 0.63 (0.51–0.73) and 0.57 (0.46–0.66), respectively, and pooled specificities were 0.88 (0.80–0.93) and 0.84 (0.77–0.89).

Platelet count was evaluated in 8 studies.16,18,19,22,24,26,30,32 A cutoff value >400 × 10⁹/L yielded lower sensitivities compared with lower
cutoff values. The pooled sensitivities with and without studies with a cutoff value <400 × 10⁹/L[18,32] were 0.55 (0.36–0.73) and 0.45 (0.28–0.63), respectively; the specificities were 0.88 (0.81–0.93) and 0.91 (0.87–0.94). The study of Beatti et al[16] was identified as an outlier in which sensitivity (0.82) was high with a cutoff of 400 × 10⁹/L. Pooled sensitivity and specificity without the data of Beatti et al were 0.37 (0.28–0.47) and 0.92 (0.87–0.95).[19,22,24,26,30]

The 9 studies evaluating hemoglobin showed that age/gender-specific cutoff values[22,23,27,30,33] had higher sensitivity than fixed cutoffs.[16,19,20,24] Pooled sensitivity increased from 0.37 (0.24–0.52) to 0.56 (0.46–0.65) when the studies with fixed cutoffs were excluded. Pooled specificity did not change: 0.90 (0.83–0.94) and 0.87 (0.77–0.93), respectively. For erythrocyte sedimentation rate (ESR) and albumin, the large variation in sensitivity could not be explained by differences in study design or cutoff value. There were no outliers.

**Fecal Markers**

One study[26] reported the diagnostic accuracy of fecal S100A12 (both sensitivity and specificity: 0.97 [0.83–1.00]). Ten studies investigating FCAl (cutoff range: 50–100 μg/g) had high sensitivities (>0.86) with small CIs (Supplemental Figure 3C). Only 3 studies reported false-negative test results.[15,24,30] The specificity in the 2 case-control studies[30,31] was lower compared with the cohort studies. FCAl exhibited a pooled sensitivity and specificity of 0.99 (0.92–1.00) and 0.65 (0.54–0.74) (Table 2), and 1.00 (0.86–1.00) and 0.69 (0.63–0.74), respectively, after exclusion of 2 case-control studies.[30,31]

**Ultrasonography**

The sensitivity of bowel wall thickness >3 mm and several other parameters measured by using ultrasonography (2 studies) in children with gastrointestinal symptoms ranged from 0.78 to 1.00, and specificity ranged from 0.55 to 0.74 (Supplemental Figure 3C).[19,28]

**Combinations of Tests**

Various studies reported on the accuracy of combinations of symptoms and/or tests, but the vast majority of combinations were only assessed in a single study (Supplemental Figure 3D-E).[16,18–20,23,25–27,30,33] Three combinations were reported in 2 studies[23,25,27,29] or 3 studies.[20,26,27] Five combinations included symptoms, 12 combinations included noninvasive tests, and 4 combinations included symptoms.
Table 2: Pooled Estimates of Diagnostic Performance of Symptoms and Noninvasive Tests for IBD in Children

<table>
<thead>
<tr>
<th>Variable</th>
<th>Studies</th>
<th>n</th>
<th>Prevalence IBD, %</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>LR Positive (95% CI)</th>
<th>LR Negative (95% CI)</th>
<th>Hypothetical Cohort With IBD Prevalence of 48%</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Reduction of Endoscopies</td>
</tr>
<tr>
<td>Abdominal pain</td>
<td>6</td>
<td>684</td>
<td>43 (19–62)</td>
<td>0.82 (0.66–0.93)</td>
<td>0.17 (0.12–0.26)</td>
<td>1.6 (0.9–1.1)</td>
<td>1.05 (0.60–1.77)</td>
<td>18</td>
</tr>
<tr>
<td>Diarrhea</td>
<td>6</td>
<td>684</td>
<td>43 (19–62)</td>
<td>0.76 (0.64–0.85)</td>
<td>0.57 (0.44–0.69)</td>
<td>1.8 (1.4–2.5)</td>
<td>0.42 (0.29–0.59)</td>
<td>42</td>
</tr>
<tr>
<td>Rectal bleeding</td>
<td>7</td>
<td>1280</td>
<td>42 (19–62)</td>
<td>0.57 (0.47–0.66)</td>
<td>0.78 (0.65–0.88)</td>
<td>2.6 (1.7–4.0)</td>
<td>0.93 (0.47–0.65)</td>
<td>64</td>
</tr>
<tr>
<td>Weight loss</td>
<td>6</td>
<td>1173</td>
<td>39 (19–60)</td>
<td>0.48 (0.31–0.65)</td>
<td>0.69 (0.55–0.81)</td>
<td>1.1 (1.1–2.3)</td>
<td>0.74 (0.56–0.93)</td>
<td>61</td>
</tr>
<tr>
<td>Noninvasive tests</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CRP</td>
<td>9</td>
<td>1146</td>
<td>49 (36–62)</td>
<td>0.63 (0.51–0.73)</td>
<td>0.88 (0.80–0.95)</td>
<td>5.1 (2.8–9.4)</td>
<td>0.42 (0.30–0.59)</td>
<td>64</td>
</tr>
<tr>
<td>ESR</td>
<td>11</td>
<td>1424</td>
<td>55 (38–67)</td>
<td>0.60 (0.53–0.67)</td>
<td>0.84 (0.80–0.89)</td>
<td>4.2 (3.3–5.3)</td>
<td>0.41 (0.33–0.50)</td>
<td>60</td>
</tr>
<tr>
<td>Platelet count</td>
<td>8</td>
<td>732</td>
<td>58 (43–62)</td>
<td>0.55 (0.36–0.73)</td>
<td>0.88 (0.80–0.93)</td>
<td>4.7 (3.1–7.1)</td>
<td>0.51 (0.34–0.70)</td>
<td>68</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>9</td>
<td>1454</td>
<td>50 (36–62)</td>
<td>0.37 (0.24–0.52)</td>
<td>0.90 (0.83–0.94)</td>
<td>3.7 (2.3–5.9)</td>
<td>0.70 (0.57–0.86)</td>
<td>77</td>
</tr>
<tr>
<td>Albumin</td>
<td>8</td>
<td>721</td>
<td>52 (42–64)</td>
<td>0.49 (0.31–0.66)</td>
<td>0.89 (0.80–0.95)</td>
<td>1.5 (1.5–2.0)</td>
<td>0.52 (0.34–0.76)</td>
<td>74</td>
</tr>
<tr>
<td>FCal</td>
<td>10</td>
<td>867</td>
<td>53 (32–82)</td>
<td>0.99 (0.92–1.00)</td>
<td>0.65 (0.54–0.74)</td>
<td>2.8 (2.1–3.7)</td>
<td>0.01 (0.00–0.13)</td>
<td>35</td>
</tr>
</tbody>
</table>

* Prevalence of IBD based on cohort studies. IBD missed indicates children with IBD and a negative index test result; IBD avoided indicates children with IBD without a positive index test result, and reduction of endoscopies indicates total number of children with a negative index test result. For calculating the latter, we assumed that in the alternative strategy all IBD patients would undergo endoscopy. LR, likelihood ratio.

**DISCUSSION**

This systematic review included 19 studies reporting on the diagnostic accuracy of symptoms, noninvasive tests, and test combinations for IBD in children. All studies were performed in referred children. The prevalence of IBD ranged from 19% to 82%, and weights ranging from 0.48 to 0.92. These findings suggest that pediatricians consider an endoscopy in these selected children without IBD with chronic gastrointestinal symptoms, and symptoms caused by IBD and noninvasive tests. The 2 or more combinations of positive tests would have missed 16 (ESR) to 30 (hemoglobin) children with IBD, and the total number of endoscopies would be reduced by 60% (ESR) to 77% (hemoglobin).
was moderate. Three studies used follow-up as the reference standard in children at low risk of IBD instead of performing an endoscopy,25,27,33 which may lead to missed diagnoses of IBD and overestimation or underestimation of test characteristics. However, no effect on test characteristics was identified by studies that used 2 reference standards. Case-control design did not influence sensitivity or specificity of symptoms or blood markers, probably because we did not include case-control studies with known IBD case subjects or healthy control subjects, and these factors are acknowledged to overestimate diagnostic accuracy.34,35

**Blood Markers**

Heterogeneity in reported sensitivities was high compared with reported specificities. We could explain part of this heterogeneity by the use of different cutoff values. Lower cutoffs for platelet count (<400 × 10^9/L) and age/gender-specific cutoffs for hemoglobin increased the sensitivity of both tests. These cutoffs are more appropriate in the triage of IBD, in which high sensitivity is important.

CRP exhibited the best overall performance of all the blood markers. This finding remained true after excluding 2 studies with extreme high sensitivity and high specificity.16,29 A narrative review evaluating biomarkers in children and adults also suggests that CRP is the best blood marker to differentiate IBD from functional gastrointestinal disorders.36 Although CRP exhibited high specificity for IBD, normal CRP levels do not exclude IBD in referred children because the sensitivity is moderate.

**Fecal Markers**

FCal had a high pooled sensitivity (0.99) and a modest pooled specificity (0.65) and is therefore a useful test to rule out IBD in children whose pediatrician suspects IBD and considers endoscopy. Only 3 of 10 studies reported false-negative results for FCal. These false-negative results might be due to the higher threshold of >100 μg/g instead of >50 μg/g used in 1 study.15 Moreover, it is possible that in these studies, FCal was measured at an earlier stage of disease compared with other studies.24,30 Children with IBD at an early stage may not yet have developed a sufficiently elevated calprotectin level. Whether the lower specificity in the case-control studies compared with the cohort studies is associated with the selection of patients remains unclear.

The diagnostic accuracy of FCal for pediatric IBD was evaluated in 4 meta-analyses,7,8,37,38 1 of which was an individual patient data analysis.38 One of the meta-analyses had methodologic limitations and included known IBD case subjects and healthy control subjects, which may lead to overestimation of the diagnostic accuracy.37 The 3 high-quality reviews reported the diagnostic accuracy of FCal in children with symptoms suggestive of IBD (average prevalence ranged from 54%–61%). The pooled sensitivity and specificity reported in the 3 reviews varied between 0.92 and 0.98 and 0.68 and 0.76, respectively, which are comparable to our results.7,8,38 The small differences might be due to inclusion of different studies and variations in the 2 × 2 tables. Van Rheenen et al8 included 2 studies in which few children were already diagnosed with IBD during feces sampling,19,40 Henderson et al17 excluded studies in which follow-up was used as a reference standard.27 and Degraeuwe et al38 included 1 study in which some of the children had known IBD during feces sampling.30 In our review, studies were excluded that included children with known IBD. Moreover, we included 1 study.31

**Urinary Markers**

Urinary markers were rarely studied and showed low discriminating power. They provided no added value in combination with other markers.

**Ultrasonography**

Ultrasonography might be a feasible test: it is noninvasive, easy accessible, and does not involve radiation. The 2 studies in our review produced different results; sensitivity ranged from 0.78 to 1.00 and specificity ranged from 0.55 to 0.74. The high specificity of 1.00 is questionable, because only 4 of the children studied did not have IBD.28 A previous systematic review regarding imaging in children with IBD recommended that ultrasonography not be used for the initial diagnosis of pediatric IBD because of its low accuracy and high interoperator variability.41 However, only 1 of the 3 included studies evaluated children, and a few children were already diagnosed with IBD when they were included. Two meta-analyses showed that the diagnostic performance of ultrasonography in adults was good; sensitivity ranged from 0.73 to 0.90, and specificity in both reviews was 0.95.42,43 Before recommending ultrasonography as a triage test for IBD in children, more studies of adequate methodologic quality are needed.

**Test Combinations**

Many different test combinations were evaluated, often only in a single study, which hampers comparison of results. Overall, the specificites of combination of tests were good, whereas sensitivities were less high and heterogeneous. The combined noninvasive test using...
“or” instead of “and” showed higher sensitivities, which is important for safely excluding IBD. The combined noninvasive test combinations with the highest sensitivity of 0.97 were the combinations “FCal or albumin” and “hemoglobin or ESR or albumin or platelets count or CRP.”

Furthermore, the added value of a test to symptoms was rarely studied. One study showed that FCal had added value to the “clinical eye” of the pediatrician. In the latter study, it is unclear how this “clinical eye” incorporated symptoms and blood markers. To investigate the optimal sequential strategy, multivariable logistic regression analyses might be used. A recently published individual patient data analysis constructed a model to predict the probability of having IBD based on FCal and the child’s age. The model correctly classified 85.5% of the children, with a sensitivity of 0.81 and a specificity of 0.92 (area under the curve: 0.92). Important predictors such as symptoms, signs, and other noninvasive tests were not included in this model. Studies or more advanced individual patient data meta-analyses are required to investigate the optimal test strategy for IBD in children with gastrointestinal symptoms. Because of varying IBD prevalence and thresholds for further testing, such strategies might differ between nonreferred and referred children.

Strengths and Limitations

The strength of the present review is that we evaluated the diagnostic accuracy of all noninvasive tests for IBD in children with gastrointestinal symptoms. Before starting the review, we discussed which noninvasive tests can be reasonably deployed in primary care. The tests should be easy to perform, rapid, and applicable in primary care. We therefore excluded tests such as MRIs, computed tomography scans, positron emission tomography, scintigraphy, barium follow-through, and serology (eg, ASCA, pANCA). Although ASCA and pANCA are simple and noninvasive, these tests often produce false-negative results and are therefore not recommended for the triage of IBD. They might be helpful in differentiating between Crohn’s disease and ulcerative colitis in children with IBD and should be reserved for specialist settings. A promising fecal marker is fecal lactoferrin; however, this marker was not included in our review because it is only studied in children with known IBD and healthy control subjects. Studies in children with symptoms suggestive of IBD are needed.

Despite the extensive search of Medline and Embase, we identified 4 publications by hand-searching the references of included publications, reviews, and guidelines. This outcome might be due to the search strategy for diagnostic accuracy studies, because these search strategies are not 100% accurate in detecting relevant studies. We chose a pragmatic approach, as the search strategy significantly reduced the number of identified studies. By hand-searching the references, we believe that all relevant studies were included. In addition, we contacted authors about incorrect or insufficient 2 × 2 tables, and this follow-up enabled the construction of optimal 2 × 2 tables of tests.

Clinical Implications

In referred children with symptoms suggestive of IBD in whom the pediatrician considers endoscopy, FCal was found to be a sensitive test for safely excluding IBD. Assuming that these children otherwise would have undergone an endoscopy, FCal would reduce the number of endoscopies by 35% at the cost of 1 missed patient with IBD. By testing for FCal, 18% of the patients without IBD would undergo an invasive procedure because of a false-positive test result. One might consider that missing a child with IBD at this level of care, is unacceptable. Therefore, a sequential strategy of tests might be more adequate. In referred children with a positive FCal test result, CRP or albumin testing could be added because of their low false-positive rate and consequent reduction of unnecessary endoscopies. However, the predictive value may change when tests are applied sequentially instead of being used in isolation.

Future research is needed to investigate sequential strategies. Children included in the studies of this systematic review were all referred children. In 16 of the 19 studies, all children underwent an endoscopy. The setting (moderate/high versus high) or the prevalence (range: 19%–82%) did not influence the sensitivity or specificity of the symptoms or tests. Therefore, the results of this systematic review are generalizable to pediatricians or pediatric gastroenterologists who evaluate children in whom they consider endoscopy to be indicated. The patient population of a pediatrician varies between different health care systems.

In health care systems in the Netherlands, United Kingdom, Scandinavia, Canada, New Zealand, and Australia, children can only be seen by a pediatrician or pediatric gastroenterologist if they are referred by a primary care physician. In the United States children can visit their general pediatrician directly. By interpreting the results of our review, one must take generalizability to the intended population into account.

An important result is that none of the studies was performed in nonreferred low-risk children. In 2005, a technical report on chronic abdominal pain in children stated that symptoms were not evaluated in nonreferred children, and blood markers were rarely studied and
only in referred children.\textsuperscript{50} Although there are now sufficient number of studies investigating noninvasive tests, it is remarkable that studies in nonreferred children are still lacking. Therefore, we could not compare the diagnostic accuracy between nonreferred and referred children. Moreover, it is impossible to extrapolate our results to populations of nonreferred children. Studies evaluating the accuracy of these tests in nonreferred children are urgently needed.

CONCLUSIONS

The present review provides an overview of symptoms, signs, and noninvasive tests for IBD in children presenting with symptoms suggestive of IBD in whom a pediatrician considers endoscopy to be indicated. In these children, symptoms alone are insufficient in triage for IBD. FCal, CRP, and albumin are of clinical value, given their ability to select children at low risk (negative FCal test result) or high risk (positive albumin or CRP test result) for IBD. Further research should investigate the accuracy of sequential testing strategies and the added values of tests beyond signs and symptoms focusing on FCal, CRP, and albumin. Before tests or a diagnostic strategy can be recommended in nonreferred, low-risk children, high-quality studies are needed in this setting.

ACKNOWLEDGMENTS

The authors thank Sjoukje van der Werf (medical information specialist, central medical library, University Medical Center Groningen) for assisting with the literature search.

ABBREVIATIONS

ASCA: anti–\textit{Saccharomyces cerevisiae} antibodies
CT: computed tomography
CI: confidence interval
CRP: C-reactive protein
ESR: erythrocyte sedimentation rate
FCal: fecal calprotectin
IBD: inflammatory bowel disease
pANCA: perinuclear antineutrophil cytoplasmic antibodies

REFERENCES

8. van Rheezen PF, Van de Vijver E, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. \textit{BMJ}. 2010;341:c3369


25. Sabery N, Bass D. Use of serologic markers as a screening tool in inflammatory bowel disease compared with elevated erythrocyte sedimentation rate and anemia. Pediatrics. 2007;119(1) Available at: www.pediatrics.org/cgi/content/full/119/1/e193


Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis
Gea A. Holtman, Yvonne Lisman-van Leeuwen, Johannes B. Reitsma and Marjolein Y. Berger

Pediatrics 2016;137:: originally published online December 17, 2015; DOI: 10.1542/peds.2015-2126

Updated Information & Services
including high resolution figures, can be found at:
/content/137/1/e20152126.full

Supplementary Material
Supplementary material can be found at:
/content/suppl/2015/12/16/peds.2015-2126.DCSupplemental.html

References
This article cites 49 articles, 10 of which can be accessed free at:
/content/137/1/e20152126.full.html#ref-list-1

Citations
This article has been cited by 1 HighWire-hosted articles:
/content/137/1/e20152126.full.html#related-urls

Subspecialty Collections
This article, along with others on similar topics, appears in the following collection(s):
Gastroenterology
/cgi/collection/gastroenterology_sub

Permissions & Licensing
Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:
/site/misc/Permissions.xhtml

Reprints
Information about ordering reprints can be found online:
/site/misc/reprints.xhtml
Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis
Gea A. Holtman, Yvonne Lisman-van Leeuwen, Johannes B. Reitsma and Marjolein Y. Berger

_Pediatrics_ 2016;137;; originally published online December 17, 2015;
DOI: 10.1542/peds.2015-2126

The online version of this article, along with updated information and services, is located on the World Wide Web at:
/content/137/1/e20152126.full