Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis

Gea A. Holtman, MSc,[a] Yvonne Lisman-van Leeuwen, PhD,[a] Johannes B. Reitsma, MD, PhD,[b] Marjolein Y. Berger, MD, PhD[a]

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.

[a]*Department of General Practice, University of Groningen, University Medical Center Groningen, Groningen, the Netherlands; and* [b]*Julius Center for Health Sciences and Primary Care, University Medical Center Utrecht, Utrecht, the Netherlands*

Ms Holtman conceptualized and designed the study, selected studies, extracted data, assessed study quality, conducted the analyses, and drafted the initial manuscript; Dr Lisman-van Leeuwen conceptualized and designed the study, selected studies, extracted data, assessed study quality, and critically reviewed and revised the manuscript; Dr Reitsma conducted the analyses and critically reviewed the manuscript; Dr Berger conceptualized and designed the study, and critically reviewed and revised the manuscript; and all authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

DOI: 10.1542/peds.2015-2126

Accepted for publication Oct 26, 2015

Address correspondence to Marjolein Y. Berger, MD, PhD, Department of General Practice, FA21, University of Groningen, University Medical Center Groningen, PO Box 196, 9700 AD, Groningen, the Netherlands. E-mail: m.y.berger@umcg.nl

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

To cite: Holtman GA, Lisman-van Leeuwen Y, Reitsma JB, et al. Noninvasive Tests for Inflammatory Bowel Disease: A Meta-analysis. *Pediatrics.* 2016;137(1):e20152126

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),[1] although increasing.[2] 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.[3] 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.[4] In contrast, an attitude of "wait and see" may cause unnecessary concerns and loss of well-being in children with IBD.[5]

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.[6] 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.[7–9] 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.[10] 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,[7–9] and guidelines on pediatric IBD.[4,11] 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. Disagreement 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).[12] 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

Domain 1: Patient selection
Risk of bias
Was a consecutive or random sample of patients enrolled?
Score "yes" if the following words are stated in the article: consecutive, random, or all patients were included between a defined time period.
Was a case-control design avoided?
Score "yes" if a case-control design was avoided.
Did the study avoid inappropriate exclusions?
Score "yes" if studies excluded no other known somatic bowel disorders than IBD or excluded IBD-unclassified.
Applicability concerns
Are there concerns that the included patients and setting do not match the topic of our study?
Score "low" if the patients were less than 18 years, had symptoms suggestive of IBD and had no previous diagnosis of IBD.
Domain 2: Index test
Risk of bias
Were the index test results interpreted without knowledge of the results of the reference standard?
Score "yes" if the index test results were interpreted without knowledge of the reference standard.
If a threshold was used, was it pre-specified?
Score "yes" if the threshold was pre-specified.
Applicability concerns
Are there concerns that the index test, its conduct, or interpretation differ from the topic of our study?
Score "low" if the study provides a clear description of the index test and definition of a positive test result.
Domain 3: Reference standard
Risk of bias
Is the reference standard likely to correctly classify the target condition?
Score "yes" if the reference test is an endoscopy of the upper and lower gastrointestinal tract, including an ileum intubation, histopathology, and/or a clinical follow-up of at least 1 year. (no reference standard is 100% sensitive or specific, but the Porto Criteria state that the ileocolonoscopy including histology is essential)
Were the reference standard results interpreted without knowledge of the results of the index test?
Score "yes" if the reference standard results were interpreted without knowledge of the results of the index test.
Applicability concerns
Are there concerns that the target condition as defined by the reference standard does not match the review question?
Score "low" if the target condition IBD was defined by using an endoscopy or clinical follow-up.
Domain 4: Flow and timing
Risk of bias
Was there an appropriate interval between index test and reference standard?
Score "yes" if the time period was 1 month or less. It is unlikely that the mucosal inflammation spontaneously disappears in one month.
Did all patients receive a reference standard?
Score "yes" if all patients receive a reference standard.
Did all patients receive the same reference standard?
Score "yes" if all patients receive the same reference standard.
Were all patients included in the analysis?
Score "yes" if all patients included in the study were included in the analysis.
If all answers of signaling questions concerning a domain were "yes", the risk of bias was judged as low. If any signalling question was answered "no" the risk of bias was judged as high. When insufficient data was reported the specific item was classified as "unclear".

FIGURE 1
Signaling Questions for Scoring the QUADAS 2

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.[13,14] 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 studies[15–28] and 5 case-control studies[29–33] are presented in Supplemental Tables 3 and 4. We requested data of incorrect[19] or insufficient[18,20,22,24–28,33] 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 risk[23,25,27,33]; in 13 studies, they were at high risk[7,15–22,26,28,31,32]; 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.

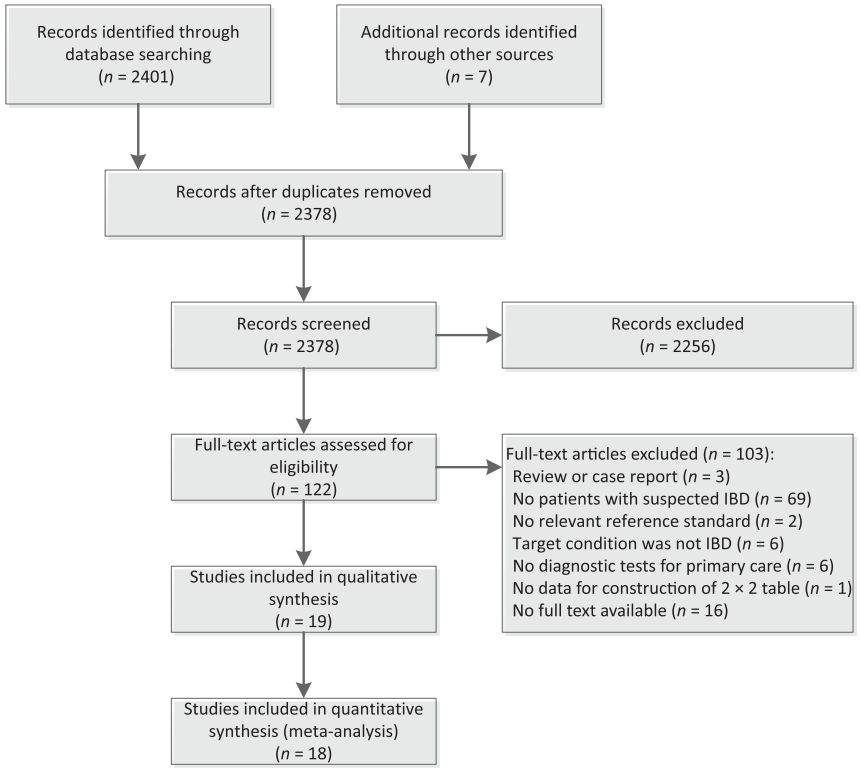

FIGURE 2
Flow diagram (according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses statement) exhibiting the elimination process for study analysis. Databases used for literature searching were Medline (*n* = 1619) and Embase (*n* = 782). Four included articles were identified by reference checking.

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 × 10^9/L yielded lower sensitivities compared with lower

TABLE 1 Summary of the Methodologic Assessment of Included Studies

Study	Risk of Bias				Applicability Concerns		
	Patient Selection	Index Test	Reference Standard	Flow and Timing	Patient Selection	Index Test	Reference Standard
Cohort studies							
Ashorn et al, 2009[15]	Unclear	Unclear	Unclear	High	High	Low	Low
Beattie et al, 1995[16]	Low	Low	Unclear	Unclear	Low	Low	Low
Bonnín et al, 2007[17]	Unclear	High	Unclear	Unclear	Low	Low	Low
Cabrera-Abreu et al, 2003[18]	Unclear	High	Unclear	Unclear	Low	Low	Low
Canani et al, 2006[19]	Low	Unclear	High	Unclear	Low	Low	Low
Diamanti et al, 2010[21]	Low	Low	Low	Unclear	Low	Low	Low
Dubinsky et al, 2001[20]	High	Unclear	High	Unclear	High	Low	Low
Fagerberg et al, 2005[22]	High	Unclear	Unclear	High	Low	Low	Low
Khan et al, 2002[23]	High	Low	High	High	Low	Low	Low
Perminow et al, 2009[24]	Unclear	Unclear	High	High	Low	Low	Low
Sabery and Bass, 2007[25]	High	Low	High	High	High	Low	Low
Sidler et al, 2008[26]	Unclear	High	Low	Unclear	Low	Low	Low
Van de Vijver et al, 2012[27]	Unclear	Low	High	High	Low	Low	Low
Ziech et al, 2014[28]	Low	Unclear	Unclear	High	Low	Low	Low
Case-control studies							
El Chammas, 2013[33]	High	Unclear	Unclear	High	Low	Low	Low
Henderson et al, 2012[30]	High	Low	High	High	Low	Low	Low
Leach et al, 2007[32]	High	Unclear	High	Low	Low	Low	Low
Minar et al, 2014[31]	High	Unclear	Unclear	High	Low	Low	Low
Tsampalieros et al, 2011[29]	High	Unclear	Unclear	Unclear	Low	Low	Low

High risk of bias in domain patient selection if the study had no consecutive or random sample of patients, case-control design,[29–33] or inappropriate exclusions.[20,22,23,25,29,32] High risk of bias in domain index test if the index results were interpreted with knowledge of the reference standard, or if the threshold was not prespecified.[17,18,26] High risk of bias in domain reference standard if the endoscopy did not include an ileum intubation,[19,30,32] the follow-up was <12 months,[27] or if the reference standard results were interpreted with knowledge of the index test.[20,23–25,30] High risk of bias in the domain flow and timing if the time period between the index test and reference standard was >1 month,[15,25,27,30] not all patients received a reference standard, not all patients received the same reference standard,[25,27,33] or not all patients were included in the analysis.[15,22–25,28,31] High applicability concerns for the domain patient selection if the study included children aged >18 years.[15,20,25]

cutoff values. The pooled sensitivities with and without studies with a cutoff value <400 × 10^9/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 × 10^9/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]

Urinary Markers

One study found that measurement of urinary excretion of cellobiose/mannitol with a cutoff of 0.023 had a sensitivity and specificity of 0.41 (0.22–0.61) and 0.67 (0.41–0.87), respectively (Supplemental Figure 3C).[19]

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

Variable	Studies	*n*	Prevalence IBD, % (range)[a]	Sensitivity (95% CI)	Specificity (95% CI)	LR Positive (95% CI)	LR Negative (95% CI)	Hypothetical Cohort With IBD Prevalence of 48%		
								Reduction of Endoscopies	IBD Missed	Unnecessary Endoscopies
Symptoms										
Abdominal pain	6	684	43 (19–62)	0.82 (0.66–0.92)	0.17 (0.12–0.24)	1.0 (0.9–1.1)	1.03 (0.60–1.77)	18	9	43
Diarrhea	6	684	43 (19–62)	0.76 (0.64–0.85)	0.57 (0.44–0.69)	1.8 (1.4–2.3)	0.42 (0.29–0.59)	42	12	22
Rectal bleeding	7	1280	42 (19–62)	0.57 (0.47–0.66)	0.78 (0.65–0.88)	2.6 (1.7–4.0)	0.55 (0.47–0.65)	62	21	11
Weight loss	6	1173	39 (19–60)	0.48 (0.31–0.66)	0.69 (0.55–0.81)	1.6 (1.1–2.3)	0.74 (0.56–0.99)	61	25	16
Noninvasive tests										
CRP	9	1146	49 (36–62)	0.63 (0.51–0.73)	0.88 (0.80–0.93)	5.1 (2.8–9.4)	0.42 (0.30–0.59)	64	18	6
ESR	11	1434	55 (36–67)	0.66 (0.58–0.73)	0.84 (0.80–0.88)	4.2 (3.3–5.3)	0.41 (0.33–0.50)	60	16	8
Platelet count	8	732	58 (43–62)	0.55 (0.36–0.73)	0.88 (0.81–0.93)	4.7 (2.9–7.7)	0.51 (0.34–0.76)	68	22	6
Hemoglobin	9	1454	50 (36–62)	0.37 (0.24–0.52)	0.90 (0.83–0.94)	3.7 (2.5–5.9)	0.70 (0.57–0.86)	77	30	5
Albumin	6	527	53 (43–62)	0.48 (0.31–0.66)	0.94 (0.86–0.98)	8.3 (3.7–18.7)	0.55 (0.40–0.76)	74	25	3
FCal	10	867	53 (32–82)	0.99 (0.92–1.00)	0.65 (0.54–0.74)	2.8 (2.1–3.7)	0.01 (0.00–0.13)	35	1	18

[a] Prevalence of IBD based on cohort studies. IBD missed indicates children with IBD and a negative index test result; unnecessary endoscopies indicates children without IBD with 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 100 patients would undergo endoscopy. LR, likelihood ratio.

and noninvasive tests. The 2 "or" combinations requiring 1 of the included tests to have positive results with highest sensitivity (0.97) were: (1) FCal or albumin; or (2) hemoglobin or ESR or albumin or platelet count or CRP. The "and" combination, requiring all included tests to be positive, with highest specificity (0.96) was hemoglobin and ESR.

Four studies investigated the additional value of a single test in addition to symptoms or other tests. Khan et al[23] found that the addition of hemoglobin and ESR to the finding of rectal bleeding increased sensitivity (from 0.68 to 0.86) and decreased specificity (from 0.92 to 0.59). Another study found that ESR had no additional value when added to platelet count and hemoglobin.[18] Cellobiose/mannitol in urine added no additional value to the combination of FCal, anti–*Saccharomyces cerevisiae* antibodies (ASCA), perinuclear antineutrophil cytoplasmic antibodies (pANCA), and ultrasonography.[19] Van de Vijver et al[27] investigated the additional value of FCal to the "clinical eye" of the pediatrician. FCal reduced the proportion of IBD-negative endoscopies from 38% to 32% without missing 1 child with IBD.

Clinical Impact

Table 2 presents the potential clinical impact of each index test based on their pooled estimates of test accuracy in 100 hypothetical children with gastrointestinal symptoms. Given that prevalence did not explain the heterogeneity of sensitivity and specificity in our study, we took the weighted mean prevalence as representative for the population on which our result might be extrapolated. In the hypothetical cohort with an IBD prevalence of 48%, testing with FCal would miss 1 child with IBD, 18 children without IBD would undergo an unnecessary endoscopy, and the number of children requiring endoscopy compared with a strategy in which all children would undergo endoscopy would be reduced by 35%. Testing with different blood markers would have missed 16 (ESR) to 30 (hemoglobin) children with IBD; 3 (albumin) to 8 (ESR) children without IBD would have had an endoscopy; and the total number of endoscopies would be reduced by 60% (ESR) to 77% (hemoglobin).

DISCUSSION

This systematic review included 19 studies reporting on the diagnostic accuracy of symptoms, signs, noninvasive tests, and test combinations for IBD in children with chronic gastrointestinal symptoms. All studies were performed in referred children. The prevalence of IBD ranged from 19% to 82%. Diagnostic accuracy of individual symptoms and signs was low, with pooled sensitivities for abdominal pain, diarrhea, rectal bleeding, and weight loss ranging from 0.48 to 0.82 and specificities ranging from 0.17 to 0.78. These findings suggest that in these selected children in whom a pediatrician considers an endoscopy to be indicated, individual signs and symptoms cannot distinguish symptoms caused by IBD from those caused by other conditions. Therefore, easy tests with few adverse effects (and is preferably noninvasive) seem essential in the triage for endoscopy. Use of FCal was best at decreasing the probability of IBD, with a pooled negative likelihood ratio of 0.01 (0.0–0.1). CRP and albumin had pooled positive likelihood ratios sufficiently high to indicate an increase in the probability of IBD that may be of clinical importance (positive likelihood ratios of 5.1 [2.8-9.4] and 8.3 [3.7-18.7], respectively).

Because 17 studies exhibited a high risk of bias in ≥1 domain, the quality of the diagnostic studies

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 ($<400 \times 10^9$/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[30,31] 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 al[8] included 2 studies in which few children were already diagnosed with IBD during feces sampling,[39,40] Henderson et al[7] excluded studies in which follow-up was used as a reference standard,[27] and Degraeuwe et al[38] included 1 study in which some of the children had known IBD during feces sampling.[40] In our review, studies were excluded that included children with known IBD. Moreover, we included 1 study[31] issued after publication of the latest review.[38]

Fecal s100A12 showed promising results with high sensitivity and specificity. More research is needed to evaluate the diagnostic accuracy of this marker.

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 sentitivities 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.[27] 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.[38] 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.[4] They might be helpful in differentiating between Crohn's disease and ulcerative colitis in children with IBD[4] 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.[44–46] 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.[47] 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.[48] 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.[49] 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.[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–*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

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: No external funding.

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.

REFERENCES

1. van der Zaag-Loonen HJ, Casparie M, Taminiau JA, Escher JC, Pereira RR, Derkx HH. The incidence of pediatric inflammatory bowel disease in the Netherlands: 1999-2001. *J Pediatr Gastroenterol Nutr*. 2004;38(3):302–307
2. Benchimol EI, Fortinsky KJ, Gozdyra P, Van den Heuvel M, Van Limbergen J, Griffiths AM. Epidemiology of pediatric inflammatory bowel disease: a systematic review of international trends. *Inflamm Bowel Dis*. 2011;17(1):423–439
3. Sawczenko A, Sandhu BK. Presenting features of inflammatory bowel disease in Great Britain and Ireland. *Arch Dis Child*. 2003;88(11):995–1000
4. Levine A, Koletzko S, Turner D, et al; European Society of Pediatric Gastroenterology, Hepatology, and Nutrition. ESPGHAN revised Porto criteria for the diagnosis of inflammatory bowel disease in children and adolescents. *J Pediatr Gastroenterol Nutr*. 2014;58(6):795–806
5. Kim SC, Ferry GD. Inflammatory bowel diseases in pediatric and adolescent patients: clinical, therapeutic, and psychosocial considerations. *Gastroenterology*. 2004;126(6):1550–1560
6. Bossuyt PM, Irwig L, Craig J, Glasziou P. Comparative accuracy: assessing new tests against existing diagnostic pathways. *BMJ*. 2006;332(7549):1089–1092
7. Henderson P, Anderson NH, Wilson DC. The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease: a systematic review and meta-analysis. *Am J Gastroenterol*. 2014;109(5):637–645
8. van Rheenen PF, Van de Vijver E, Fidler V. Faecal calprotectin for screening of patients with suspected inflammatory bowel disease: diagnostic meta-analysis. *BMJ*. 2010;341:c3369
9. Waugh N, Cummins E, Royle P, et al. Faecal calprotectin testing for differentiating amongst inflammatory and non-inflammatory bowel diseases: systematic review and economic evaluation. *Health Technol Assess*. 2013;17(55):xv–xix, 1–211
10. Leeflang MM, Scholten RJ, Rutjes AW, Reitsma JB, Bossuyt PM. Use of methodological search filters to identify diagnostic accuracy studies can lead to the omission of relevant studies. *J Clin Epidemiol*. 2006;59(3):234–240
11. DeRidder L, Rings EH, Escher JC. Guideline "Diagnosis and treatment of inflammatory bowel disease in children" [in Dutch]. *Ned Tijdschr Geneeskd*. 2009;154(A):1898
12. Whiting PF, Rutjes AW, Westwood ME, et al; QUADAS-2 Group. QUADAS-2: a

revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med*. 2011;155(8):529–536

13. Reitsma JB, Glas AS, Rutjes AW, Scholten RJ, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *J Clin Epidemiol*. 2005;58(10):982–990
14. Diaz M. Performance measures of the bivariate random effects model for meta-analyses of diagnostic accuracy. *Comput Stat Data Anal*. 2015;83:82–90
15. Ashorn S, Honkanen T, Kolho KL, et al. Fecal calprotectin levels and serological responses to microbial antigens among children and adolescents with inflammatory bowel disease. *Inflamm Bowel Dis*. 2009;15(2):199–205
16. Beattie RM, Walker-Smith JA, Murch SH. Indications for investigation of chronic gastrointestinal symptoms. *Arch Dis Child*. 1995;73(4):354–355
17. Bonnín Tomàs A, Vila Vidal M, Rosell Camps A. Fecal calprotectin as a biomarker to distinguish between organic and functional gastrointestinal disease [in Spanish]. *Rev Esp Enferm Dig*. 2007;99(12):689–693
18. Cabrera-Abreu JC, Davies P, Matek Z, Murphy MS. Performance of blood tests in diagnosis of inflammatory bowel disease in a specialist clinic. *Arch Dis Child*. 2004;89(1):69–71
19. Canani RB, de Horatio LT, Terrin G, et al. Combined use of noninvasive tests is useful in the initial diagnostic approach to a child with suspected inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2006;42(1):9–15
20. Dubinsky MC, Ofman JJ, Urman M, Targan SR, Seidman EG. Clinical utility of serodiagnostic testing in suspected pediatric inflammatory bowel disease. *Am J Gastroenterol*. 2001;96(3):758–765
21. Diamanti A, Panetta F, Basso MS, et al. Diagnostic work-up of inflammatory bowel disease in children: the role of calprotectin assay. *Inflamm Bowel Dis*. 2010;16(11):1926–1930
22. Fagerberg UL, Lööf L, Myrdal U, Hansson LO, Finkel Y. Colorectal inflammation is well predicted by fecal calprotectin in children with gastrointestinal symptoms. *J Pediatr Gastroenterol Nutr*. 2005;40(4):450–455
23. Khan K, Schwarzenberg SJ, Sharp H, Greenwood D, Weisdorf-Schindele S. Role of serology and routine laboratory tests in childhood inflammatory bowel disease. *Inflamm Bowel Dis*. 2002;8(5):325–329
24. Perminow G, Brackmann S, Lyckander LG, et al; IBSEN-II Group. A characterization in childhood inflammatory bowel disease, a new population-based inception cohort from South-Eastern Norway, 2005-07, showing increased incidence in Crohn's disease. *Scand J Gastroenterol*. 2009;44(4):446–456
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
26. Sidler MA, Leach ST, Day AS. Fecal S100A12 and fecal calprotectin as noninvasive markers for inflammatory bowel disease in children. *Inflamm Bowel Dis*. 2008;14(3):359–366
27. Van de Vijver E, Schreuder AB, Cnossen WR, Muller Kobold AC, van Rheenen PF; North Netherlands Pediatric IBD Consortium. Safely ruling out inflammatory bowel disease in children and teenagers without referral for endoscopy. *Arch Dis Child*. 2012;97(12):1014–1018
28. Ziech ML, Hummel TZ, Smets AM, et al. Accuracy of abdominal ultrasound and MRI for detection of Crohn disease and ulcerative colitis in children. *Pediatr Radiol*. 2014;44(11):1370–1378
29. Tsampalieros A, Griffiths AM, Barrowman N, Mack DR. Use of C-reactive protein in children with newly diagnosed inflammatory bowel disease. *J Pediatr*. 2011;159(2):340–342
30. Henderson P, Casey A, Lawrence SJ, et al. The diagnostic accuracy of fecal calprotectin during the investigation of suspected pediatric inflammatory bowel disease. *Am J Gastroenterol*. 2012;107(6):941–949
31. Minar P, Haberman Y, Jurickova I, et al. Utility of neutrophil Fcγ receptor I (CD64) index as a biomarker for mucosal inflammation in pediatric Crohn's disease. *Inflamm Bowel Dis*. 2014;20(6):1037–1048
32. Leach ST, Yang Z, Messina I, et al. Serum and mucosal S100 proteins, calprotectin (S100A8/S100A9) and S100A12, are elevated at diagnosis in children with inflammatory bowel disease. *Scand J Gastroenterol*. 2007;42(11):1321–1331
33. El-Chammas K, Majeskie A, Simpson P, Sood M, Miranda A. Red flags in children with chronic abdominal pain and Crohn's disease—a single center experience. *J Pediatr*. 2013;162(4):783–787
34. Lijmer JG, Mol BW, Heisterkamp S, et al. Empirical evidence of design-related bias in studies of diagnostic tests. *JAMA*. 1999;282(11):1061–1066
35. Rutjes AW, Reitsma JB, Di Nisio M, Smidt N, van Rijn JC, Bossuyt PM. Evidence of bias and variation in diagnostic accuracy studies. *CMAJ*. 2006;174(4):469–476
36. Vermeire S, Van Assche G, Rutgeerts P. Laboratory markers in IBD: useful, magic, or unnecessary toys? *Gut*. 2006;55(3):426–431
37. von Roon AC, Karamountzos L, Purkayastha S, et al. Diagnostic precision of fecal calprotectin for inflammatory bowel disease and colorectal malignancy. *Am J Gastroenterol*. 2007;102(4):803–813
38. Degraeuwe PL, Beld MP, Ashorn M, et al. Faecal calprotectin in suspected paediatric inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2015;60(3):339–346
39. Bunn SK, Bisset WM, Main MJ, Gray ES, Olson S, Golden BE. Fecal calprotectin: validation as a noninvasive measure of bowel inflammation in childhood inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2001;33(1):14–22
40. Kolho KL, Raivio T, Lindahl H, Savilahti E. Fecal calprotectin remains high during glucocorticoid therapy in children with inflammatory bowel disease. *Scand J Gastroenterol*. 2006;41(6):720–725
41. Duigenan S, Gee MS. Imaging of pediatric patients with inflammatory bowel disease. *AJR Am J Roentgenol*. 2012;199(4):907–915

42. Jellema P, van Tulder MW, van der Horst HE, Florie J, Mulder CJ, van der Windt DA. Inflammatory bowel disease: a systematic review on the value of diagnostic testing in primary care. *Colorectal Dis*. 2011;13(3):239–254

43. Horsthuis K, Bipat S, Bennink RJ, Stoker J. Inflammatory bowel disease diagnosed with US, MR, scintigraphy, and CT: meta-analysis of prospective studies. *Radiology*. 2008;247(1):64–79

44. Walker TR, Land ML, Kartashov A, et al. Fecal lactoferrin is a sensitive and specific marker of disease activity in children and young adults with inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2007;44(4):414–422

45. Pfefferkorn MD, Boone JH, Nguyen JT, Juliar BE, Davis MA, Parker KK. Utility of fecal lactoferrin in identifying Crohn disease activity in children. *J Pediatr Gastroenterol Nutr*. 2010;51(4):425–428

46. Joishy M, Davies I, Ahmed M, et al. Fecal calprotectin and lactoferrin as noninvasive markers of pediatric inflammatory bowel disease. *J Pediatr Gastroenterol Nutr*. 2009;48(1):48–54

47. Whiting P, Westwood M, Beynon R, Burke M, Sterne JA, Glanville J. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. *J Clin Epidemiol*. 2011;64(6):602–607

48. Moons KG, Biesheuvel CJ, Grobbee DE. Test research versus diagnostic research. *Clin Chem*. 2004;50(3):473–476

49. Thomson S, Osborn R, Squires D, Reed SJ. International profiles of health care systems 2011: Australia, Canada, Denmark, England, France, Germany, Iceland, Italy, Japan, the Netherlands, New Zealand, Norway, Sweden, Switzerland, and the United States. 2011

50. Di Lorenzo C, Colletti RB, Lehmann HP, et al; AAP Subcommittee; NASPGHAN Committee on Chronic Abdominal Pain. Chronic abdominal pain in children: a technical report of the American Academy of Pediatrics and the North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. *J Pediatr Gastroenterol Nutr*. 2005;40(3):249–261

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

PEDIATRICS®
OFFICIAL JOURNAL OF THE AMERICAN ACADEMY OF PEDIATRICS

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