Diagnostic strategies in children with chronic gastrointestinal symptoms in primary care
Holtman, Geeske

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2016

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
CHAPTER 5

DIAGNOSTIC ACCURACY OF FAecal CALPROTECTIN FOR PAEDIATRIC INFLAMMATORY BOWEL DISEASE IN PRIMARY CARE, A PROSPECTIVE COHORT STUDY

Gea A Holtman
Yvonne Lisman-van Leeuwen
Boudewijn J Kollen
Obbe F Norbruis
Johanna C Escher
Angelika Kindermann
Yolanda B de Rijke
Patrick F van Rheenen
Marjolein Y Berger

Annals of Family Medicine; in press
ABSTRACT

PURPOSE
In specialist care, faecal calprotectin (FCal) is a commonly used non-invasive diagnostic test to rule out inflammatory bowel disease (IBD) in children with chronic gastrointestinal symptoms. The aim of this study was to evaluate the diagnostic accuracy of FCal for IBD in symptomatic children in primary care.

METHODS
We studied 2 prospective cohorts of children with chronic diarrhoea, recurrent abdominal pain, or both: children initially seen in primary care (primary care cohort) and children referred to specialist care (referred cohort). FCal (index test) was measured at baseline and compared with 1 of 2 reference standards for IBD: endoscopic assessment or 1-year follow-up. Physicians were blinded to FCal results, and values greater than 50 μg/g faeces were considered positive. We determined specificity in primary care cohort and sensitivity in referred cohort.

RESULTS
None of the 114 children included in the primary care cohort ultimately received diagnosis of IBD. The specificity of FCal in the primary care cohort was 0.87 (95% CI, 0.80-0.92). Among the 90 children in the referred cohort, 17 patients (19%) ultimately received a diagnosis of IBD. The sensitivity of FCal in the referred cohort was 0.99 (95% CI, 0.81-1.00).

CONCLUSIONS
The findings of this study suggest that a positive FCal result in children with chronic gastrointestinal symptoms seen in primary care is not likely to be indicative of IBD. However, a negative FCal result is likely to be a true negative, which safely rules out IBD in children in whom a general practitioner considers referral to specialist care.

INTRODUCTION
General practitioners (GPs) frequently manage recurrent abdominal pain or diarrhoea in children. These symptoms account for approximately 2% to 5% of all childhood consultations. Although they are typically functional in origin, it is essential that organic disease be ruled out. An organic disease that GPs should not miss is inflammatory bowel disease (IBD), that is, Crohn disease and ulcerative colitis. Delay in diagnosing IBD, and the resultant delay in receipt of appropriate treatment, may prolong suffering and can lead to complications such as anaemia, irreversible growth failure, and delayed sexual maturation.

According to guidelines, GPs should refer children with chronic diarrhoea, recurrent abdominal pain, or both for specialist care if red flags are present. However, the red flags are nonspecific and discriminate poorly between functional and organic gastrointestinal diseases, often leading to referral and extensive diagnostic testing. For children with functional disorders, referral or extensive testing may delay appropriate interventions and further decrease well-being.

Calprotectin is a calcium-binding protein released from neutrophils during intestinal inflammation that can be easily measured in faeces. In specialist care, evidence shows it to be a useful, simple, non-invasive test that can rule out IBD in children with gastrointestinal symptoms. However, the diagnostic accuracy of faecal calprotectin (FCal) has not been assessed in children evaluated in primary care. Primary and specialist care often have different populations, case mixes, and disease severity, which can affect the pre-test probability of IBD and the sensitivity and specificity of the FCal test. The diagnostic accuracy of FCal in the primary care setting should therefore be clarified before this test is recommended for routine use in primary care. We set out to study the diagnostic accuracy of FCal for identifying IBD in children with chronic gastrointestinal symptoms in primary care.

METHODS

STUDY DESIGN
This was a prospective cohort study with a delayed-type cross-sectional design. Children in the Netherlands with chronic gastrointestinal symptoms were included from July 2011 to July 2013 and had 12 months of follow-up. We studied 2 cohorts: 1) the primary care cohort consisted of consecutive children who were seen by any of 64 GPs (38 practices); 2) the referred cohort consisted of consecutive children who were referred for diagnostic work-up by GPs and general paediatricians to any of 4 general hospitals or 3 academic centres, as well as children selected from the primary care cohort based on the presence of at least 1 red flag (Figure 1). The medical ethics committee of the University Medical Centre Groningen approved the study. Written informed consent was provided by the parents of all children and by all children aged 12 years or older. The study design has been described in more detail elsewhere.

PARTICIPANTS
Children aged 4 to 18 years who sought care for chronic diarrhoea, recurrent abdominal pain,
Chronic diarrhoea was defined as soft to watery stool (score of 5, 6, or 7 on the Bristol stool chart) for at least 2 weeks or at least 2 episodes in the past 6 months. Recurrent abdominal pain was defined as at least 2 episodes of abdominal pain or discomfort in the past 6 months. Children were excluded if they had a previous diagnosis of chronic organic gastrointestinal disease; an evaluation with endoscopy or FCal for gastrointestinal symptoms in 6 months before this study; or difficulty in understanding questionnaires. Furthermore, we excluded children with long-term use (>3 months) of antibiotics, non-steroid anti-inflammatory drugs, or oral corticosteroids in the past 6 months, as well as those aged younger than 4 years, because previous studies have demonstrated elevated calprotectin concentrations in these groups. 23,24

BASELINE EVALUATION

All participating physicians assessed children for the presence of 10 red flags suggestive of IBD using a structured evaluation form. These red flags consisted of 6 alarm symptoms and 4 blood markers (Table 1). Faeces were tested for pathogens – Salmonella enterica, Campylobacter

**Table 1. Definitions for red flags of inflammatory bowel disease.**

<table>
<thead>
<tr>
<th>Red flag</th>
<th>Method of Ascertainment</th>
<th>Definition of positive finding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Involuntary weight loss</td>
<td>History</td>
<td>Involuntary decrease in weight of &gt;1 kg</td>
</tr>
<tr>
<td>Rectal blood loss</td>
<td>History</td>
<td>Rectal blood loss with defecation without constipation according to ROME III criteria</td>
</tr>
<tr>
<td>Family history of IBD</td>
<td>History</td>
<td>First-degree relatives</td>
</tr>
<tr>
<td>Growth failure</td>
<td>History and physical examination</td>
<td>Target height range &gt; -1 standard deviation score</td>
</tr>
<tr>
<td>Extra-intestinal symptoms</td>
<td>Physical examination</td>
<td>Eyes (episcleritis, scleritis, uveitis), skin (erythema nodosum, pyoderma gangrenosum, psoriasis), mouth ulcers, finger clubbing, arthritis</td>
</tr>
<tr>
<td>Peri-anal lesions</td>
<td>Physical examination</td>
<td>skin tags, hemorrhoids, fissures, fistulas, abscess</td>
</tr>
<tr>
<td>Blood markers</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haemoglobin</td>
<td>Local laboratory</td>
<td>4-12 years &lt;7.1 mmol/L, boys 12-18 years &lt;8.1 mmol/L, girls 12-18 years &lt;7.4 mmol/L</td>
</tr>
<tr>
<td>C-reactive protein</td>
<td>Local laboratory</td>
<td>&gt;10 mg/L</td>
</tr>
<tr>
<td>Erythrocyte sedimentation rate</td>
<td>Local laboratory</td>
<td>&gt;20 mm/h</td>
</tr>
<tr>
<td>Platelet count</td>
<td>Local laboratory</td>
<td>&gt;450 ×10^9/L</td>
</tr>
</tbody>
</table>

Abbreviations: IBD, inflammatory bowel disease.
Table 2. Baseline characteristics of primary care cohort and referred cohort.

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Main analysis</th>
<th>Referred cohort by origin</th>
<th>Referred cohort by referral</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Primary care</td>
<td>Referred by GP 65%</td>
<td>Referred by general</td>
</tr>
<tr>
<td></td>
<td>cohort (n = 114)</td>
<td>(n = 90)</td>
<td>paediatrician (n = 23)</td>
</tr>
<tr>
<td>Male (n %)</td>
<td>38 (33)</td>
<td>37 (41)</td>
<td>29 (44)</td>
</tr>
<tr>
<td>Age at baseline (median, IQR)</td>
<td>9 (6-12)</td>
<td>11 (7-15)</td>
<td>12 (12-15)</td>
</tr>
<tr>
<td>Presenting symptoms (n %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Recurrent abdominal pain</td>
<td>8 (8.7)</td>
<td>7 (7.8)</td>
<td>9 (9.3)</td>
</tr>
<tr>
<td>- Inflammatory bowel disease</td>
<td>6 (5.6)</td>
<td>6 (6.7)</td>
<td>6 (6.4)</td>
</tr>
<tr>
<td>- Family history of IBD</td>
<td>5 (5.4)</td>
<td>11 (12.2)</td>
<td>6 (6.3)</td>
</tr>
<tr>
<td>- Growth failure</td>
<td>4 (3.5)</td>
<td>6 (6.7)</td>
<td>3 (3.3)</td>
</tr>
<tr>
<td>- Extra-intestinal symptoms</td>
<td>0 (0)</td>
<td>13 (14.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>- Per-rectal lesions</td>
<td>7 (6.5)</td>
<td>13 (14)</td>
<td>7 (7.2)</td>
</tr>
<tr>
<td>Positive Blood markers (n/N %)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>- Haemoglobin ≥ 11.2 mmol/L</td>
<td>20 (18)</td>
<td>11 (12)</td>
<td>19 (26)</td>
</tr>
<tr>
<td>- CRP (n = 111) ≤ 10 mg/L</td>
<td>41 (36)</td>
<td>30 (34)</td>
<td>32 (44)</td>
</tr>
<tr>
<td>- Anti-tissue transglutaminase ≤ 250</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
</tbody>
</table>

Abbreviations: CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; GP, general practitioner; IBD, inflammatory bowel disease.

1 Five children with red flags were found in a patient cohort and 20 children in the referred cohort.
2 Age and sex specific: female sex: <7.1 mmol/L; aged 12-18 years: boys <8.1 mmol/L, girls <7.4 mmol/L.
3 Twenty-five children had IgA deficiency.
4 Recurrent abdominal pain, extra-intestinal symptoms, per-rectal lesions, positive blood markers (Hb, CRP, ESR, platelet count).
Given an expected sensitivity of 95%, a 95% CI and precision of ±10%, an IBD prevalence of 20%, and a loss to follow-up of 10%, we needed to include 100 children in the referred cohort.

For secondary outcomes, we calculated specificity, post-test probability, and area under the receiver operating characteristic curve (AUC) with 95% CI in the referred cohort. We also determined the effect of different FCal cut-off values (>50 μg/g, >100 μg/g, >250 μg/g) on the test characteristics, number of referrals, and missed diagnoses of IBD.

We assumed that children included in specialist care were comparable to those with at least 1 red flag from the primary care cohort. To test this assumption, we compared the characteristics of these groups. In order to evaluate spectrum bias (whereby the test setting affects the test performance), we performed subgroup analyses in the referred cohort.

Table 3. Prevalence of symptoms, blood marker positivity, and FCal positivity by final diagnosis.

<table>
<thead>
<tr>
<th>Diagnosis</th>
<th>N (%)</th>
<th>symptom positive</th>
<th>Blood marker positive</th>
<th>FCal &gt;50 μg/g</th>
<th>Range FCal (μg/g)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Primary care cohort</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional gastrointestinal disorder</td>
<td>108 (95)</td>
<td>24</td>
<td>9/104</td>
<td>12/98</td>
<td>20-257</td>
</tr>
<tr>
<td>Gastroenteritis*</td>
<td>5 (45)</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>20-88</td>
</tr>
<tr>
<td>Refused endoscopy</td>
<td>1 (1)</td>
<td>1</td>
<td>0</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Referred cohort</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBD</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crohn’s disease</td>
<td>7 (8)</td>
<td>7</td>
<td>7</td>
<td>6/6</td>
<td>152-2823</td>
</tr>
<tr>
<td>Ulcerative colitis</td>
<td>8 (9)</td>
<td>7</td>
<td>4/7</td>
<td>8</td>
<td>53-916</td>
</tr>
<tr>
<td>IBD unclassified</td>
<td>2 (2)</td>
<td>2</td>
<td>1</td>
<td>2</td>
<td>79-778</td>
</tr>
<tr>
<td><strong>Non-IBD</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional gastrointestinal disorder</td>
<td>66 (73)</td>
<td>40</td>
<td>12/87</td>
<td>10/63</td>
<td>20-185</td>
</tr>
<tr>
<td>Gastroenteritis*</td>
<td>3 (3)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20-45</td>
</tr>
<tr>
<td>Reflex esophagitis</td>
<td>1 (1)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>22</td>
</tr>
<tr>
<td>Celiac disease</td>
<td>1 (1)</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Solitary rectum ulcer</td>
<td>1 (1)</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>299</td>
</tr>
</tbody>
</table>

Abbreviations: FCal, faecal calprotectin; IBD, inflammatory bowel disease.

* Presence of 1 or more of the following: growth failure, involuntary weight loss, rectal blood loss, extra-intestinal symptoms, peri-anal lesions, family history of IBD.

+ Blood markers: Haemoglobin (4-12 years <7.1 mmol/L; 12-18 years: boys <8.1 mmol/L, girls <7.4 mmol/L), C-reactive protein (>10 mg/L), erythrocyte sedimentation rate (>20 mm/h), platelet count (>450 x10^9/L).

† Gastroenteritis due to salmonella enteric (0 cases included by GP, 2 cases included by paediatrician), Shiga Toxigenic Escherichia Coli (STEC) (1 and 0), Giardia lamblia (4 and 1).

Note: One child refused endoscopy and evaluation of red flags at 12 months’ follow-up, so the diagnosis was unknown. Nine children without IBD, including 1 child with a solitary rectal ulcer, underwent upper and lower endoscopy, including ileal intubation. The remaining 3 children did not undergo complete endoscopic evaluation for various reasons: the colonoscopy was prematurely terminated because of mucosal bleeding in 1 child with a functional gastrointestinal disorder, but was not repeated because their symptoms subsided; 1 child with a functional gastrointestinal disorder underwent colonoscopy only, but not esophagastroduodenoscopy; and 1 child received a diagnosis of celiac disease by esophagastroduodenoscopy only.

Figure 2. Flow charts and contingency tables for the calculation of diagnostic accuracy in the primary care cohort and referred cohort, using the non-imputed dataset.

Abbreviations: FC, Faecal calprotectin; GI, gastrointestinal; IBD, inflammatory bowel disease; PPV, positive predictive value; NPV, negative predictive value.

Note: The left flow chart shows the specificity of FCal (>50 μg/g) for IBD in primary care cohort (11 missing values). Specificity of standard follow-up and endoscopy was 0.88 (95% CI, 0.80–0.93) and 0.50 (95% CI, 0.09–0.91), respectively. The right flow chart shows the test characteristics of FCal (>50 μg/g) for IBD in referred cohort (5 missing values). Sensitivity of reference standard follow-up and endoscopy were 1.00 (95% CI, 0.34–1.00) and 1.00 (95% CI, 0.78–1.00), respectively; values of specificity were 0.87 (95% CI, 0.76–0.93) and 0.67 (95% CI, 0.35–0.88), respectively.

Comparing children who were referred by a GP with those referred by a general paediatrician. To evaluate the likelihood of differential verification bias, we separately calculated the test characteristics for both endoscopy and 12-month follow-up.

We conducted a missing value analysis to rule out “missing not at random” as possible explanation for missing data from the variables FCal and diagnosis. The missing data were replaced based on a multiple imputation procedure (conditional specification, predictive mean matching, 20 iterations, and 20 datasets). The patient characteristics, all symptoms, all diagnostic tests, setting, endoscopy performed, and whether IBD was diagnosed were
of the referred cohort ultimately received a diagnosis of IBD.

\[ \text{Abbreviations: CI, confidence interval; NPV, negative predictive value; PPV, positive predictive value.} \]

\* Denominator is the 90 children in the referred cohort

\* Denominator is the 17 children in the referred cohort ultimately given a diagnosis of IBD

Table 4. Test characteristics at increasing calprotectin cut-off levels in the referred cohort using the imputed dataset (n=90).

<table>
<thead>
<tr>
<th>Test characteristics</th>
<th>&gt;50 pg/g</th>
<th>&gt;100 pg/g</th>
<th>&gt;250 pg/g</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sens (95% CI)</td>
<td>0.99 (0.81–1.00)</td>
<td>0.87 (0.65–0.96)</td>
<td>0.81 (0.58–0.93)</td>
</tr>
<tr>
<td>Spec (95% CI)</td>
<td>0.84 (0.74–0.91)</td>
<td>0.93 (0.84–0.97)</td>
<td>0.98 (0.92–0.99)</td>
</tr>
<tr>
<td>PPV (95% CI)</td>
<td>0.60 (0.42–0.76)</td>
<td>0.74 (0.53–0.88)</td>
<td>0.92 (0.69–0.98)</td>
</tr>
<tr>
<td>NPV (95% CI)</td>
<td>1.00 (0.94–1.00)</td>
<td>0.97 (0.89–0.99)</td>
<td>0.96 (0.88–0.98)</td>
</tr>
<tr>
<td>Fewer referrals (n (%))</td>
<td>61 (68%)</td>
<td>69 (77%)</td>
<td>74 (82%)</td>
</tr>
<tr>
<td>Missed cases of IBD (n (%))</td>
<td>0 (9%)</td>
<td>2 (12%)</td>
<td>3 (18%)</td>
</tr>
</tbody>
</table>

Results of the analyses on the non-imputed and imputed dataset were compared to assess the effect of multiple imputations on diagnostic accuracy. Statistical analyses were performed with IBM SPSS for Windows, Version 20.0 (IBM Corp).

**RESULTS**

**PARTICIPANTS**

The primary care cohort had 114 children, and the referred cohort had 90 children (24 with red flags from primary care plus 66 children from specialist care) (Figure 1). Table 2 shows that the 66 children from specialist care more often had weight loss, extra-intestinal symptoms, and decreased haemoglobin levels compared with the 24 children with red flags from primary care. The 25 children who were referred to the paediatric gastroenterologist by a general paediatrician showed a higher IBD prevalence than 65 children who were referred by their GP (12 vs 5 cases).

**DIAGNOSES**

None of the children in the primary care cohort received a diagnosis of IBD (Table 3). The final diagnosis was based on 12 months’ follow-up in 111 children and endoscopy in 2 children. Of the 90 children in the referred cohort, 29 (32%) underwent endoscopic evaluation, and 17 (19%) ultimately received a diagnosis of IBD.

**FAECAL CALPROTECTIN**

The median intervals from stool collection to endoscopy were 4 days and 8 days for children with IBD and without IBD, respectively; however, 11 of the 27 children (2 had missing samples) who underwent endoscopy experienced a delay of more than 1 month. Two children included as predictors. We used Rubin’s rule to calculate the pooled AUC. Results of the analyses on the non-imputed and imputed dataset were compared to assess the effect of multiple imputations on diagnostic accuracy. Statistical analyses were performed with IBM SPSS for Windows, Version 20.0 (IBM Corp).

This is the first study to evaluate the test characteristics of FCal as a marker for IBD in children seen for chronic gastrointestinal symptoms in primary care. None of the children in the primary care cohort ultimately received a diagnosis of IBD, suggesting that children with chronic gastrointestinal symptoms should not be referred directly for evaluation of IBD. Current guidelines recommend referral for diagnostic work-up based on the presence of red flags. In our primary care cohort, referrals based on red flags would have resulted in a higher false-positive referral rate for IBD compared to referrals based on FCal exceeding 50 pg/g faeces. To date, however, there is insufficient evidence to recommend FCal as a tool to guide decisions about referral for diagnostic work-up of all children with chronic gastrointestinal symptoms seen in primary care.

FCal showed high sensitivity (0.99, 95% CI, 0.85–1.00) in the referred cohort. Therefore, a negative FCal may safely rule out IBD and therewith reduce the number of referrals for evaluation of IBD in children whom the GP considers a referral. Nevertheless, the 95% CIs of false-negative rates are large, because of the relatively small numbers of children with IBD included in our study, and should be interpreted with caution. In this study we focused on IBD, but from the point of view of the GP, it is important to determine whether the symptoms are related to any organic disease. In addition to 17 children with IBD, 3 children were found to have other organic diseases (1 had celiac disease, 1 had reflux esophagitis, 1 had solitary rectum ulcer). Children with celiac disease and reflux esophagitis had normal FCal levels and would have been missed if a referral had been solely based on the results of this test.

Cut-off points might need to be higher in primary care to maintain a high negative predictive value. Although we found that an increase of the threshold from 50 μg/g to 250 μg/g faeces in the referred cohort would lead to an extra 14% reduction in referrals for diagnostic work-up for IBD, but would also increase the percentage of missed IBD diagnoses from 0% to 18% (Table 4).
µg/g faeces reduced referrals by 14% (with a drop from 33% to 18% referred), this threshold also led to false-negative results and missed IBDs (with an increase from 6% to 18% cases missed) in the referred cohort. A pragmatic approach may be to monitor children with an initial calprotectin value between 50 µg/g and 250 µg/g faeces. Children whose symptoms persist and whose calprotectin values remain high can still be referred later. A similar approach has been suggested for adults in primary care, where it was suggested that patients with irritable bowel syndrome and an initial FCal value between 50 µg/g and 150 µg/g faeces who had persistent symptoms without treatment should be re-tested after 3 months.\textsuperscript{18}

The specificity of 0.87 that we identified in the primary care cohort is higher than that reported in studies performed in specialist care, where the pooled specificity ranged between 0.68 and 0.76.\textsuperscript{19-22} We expected lower specificities because the patient mix was thought to be more diverse in primary care and because calprotectin concentrations increase in conditions such as gastroenteritis.\textsuperscript{23} Moreover, in the referred cohort, the specificity of FCal was lower in children who underwent endoscopy (0.67) than in those who received clinical follow-up (0.87). The higher specificity might be explained by higher numbers of children with functional disease in the primary care cohort and in children with clinical follow-up in the referred cohort. Consequently, the test setting (primary versus specialist care) might affect the specificity of FCal (spectrum bias).

We found a sensitivity of 0.99 in the referred cohort, which is comparable to that reported in other studies performed in specialist care (IBD prevalence of about 6%), where the pooled sensitivity ranged between 0.92 and 0.98.\textsuperscript{19-22} The reported sensitivity of FCal in this study might be an overestimation of that in children in whom a GP considers a referral for diagnostic work-up. However, in the subgroup of children referred by a GP (IBD prevalence of 8%) the sensitivity was comparable to that of children referred by a general paediatrician (IBD prevalence of 48%) (0.98 vs 1.00). We therefore assume that spectrum bias did not substantially affect the sensitivity of FCal. Lack of spectrum bias might be explained by the fact that the intestinal inflammation caused by IBD is usually severe enough to increase the calprotectin value to more than 50 µg/g faeces, even in the early stages of the disease. A meta-analysis showed that sensitivity remains stable over a range of prevalences and was not substantially influenced by spectrum bias.\textsuperscript{24} A normal FCal level thus could be used to prevent a referral of children with functional symptoms to specialist care.

Ideally, a study of FCal should include consecutive children with symptoms suggestive of IBD initially evaluated in primary care. However, this design would be extremely time consuming and costly. In order to reflect current daily, real-world practice, we included a cohort of children in whom the GP considered a referral for diagnostic work-up for chronic gastrointestinal symptoms. We assumed that the children first seen in specialist care were comparable to the selected children with red flags seen in the primary care cohort. We are confident that this assumption is valid as there were only a few differences in characteristics between these groups. However, children referred by a GP had a lower probability of IBD (8%) than those referred by a general paediatrician (48%). These findings are consistent with what one might expect in the Dutch healthcare system, where children can consult a paediatrician only after obtaining a referral from their GP, and a paediatric gastroenterologist can be consulted after a referral from a GP or general paediatrician. Comparable healthcare systems exist in the United Kingdom, Scandinavia, Canada, New Zealand, and Australia.\textsuperscript{25}

Not all patients received the same reference standard test, which may cause differential verification bias.\textsuperscript{26} As it is unethical to perform endoscopy in children with a low likelihood of organic gastrointestinal disease, these children received follow-up evaluations during 1 year. An important aspect for deciding whether the verification lead to biased estimates of accuracy is the length of the follow-up period. We are confident that we identified all children with IBD, because of the extremely low probability that a child without red flags or indication for endoscopy during 1 year of follow-up has the disease.\textsuperscript{27} We did not identify new cases of IBD at the end of the follow-up period, even among children who developed red flags. Only 2 children ultimately received a diagnosis of IBD within the follow-up period (both before 6 months). The use of follow-up in children in whom endoscopy is not considered ethical does not correspond very well with the ideal situation that arise when the diagnosis of all children is determined by endoscopy. However, following children during 1 year is the best option given the reality of clinical care.\textsuperscript{28}

To recommend a test in a new setting, the diagnostic value of that test needs to be investigated in that setting.\textsuperscript{17} We found that, in selected children in whom a GP considers a referral, FCal has satisfactory discriminating power between children with and without IBD. However, of greater clinical relevance, is whether FCal can add to the diagnostic information that is readily available from a thorough history and physical examination.\textsuperscript{29} Moreover, the added value of commonly used blood markers should be compared with the added value of FCal. Further research is therefore needed to determine whether FCal should be incorporated into the routine diagnostic evaluation of paediatric patients with chronic gastrointestinal symptoms and red flags in primary care. In addition, research should be performed to evaluate the cost-effectiveness of FCal in primary care.
REFERENCES


