Diagnostic Value of Abdominal Radiography in Constipated Children

A Systematic Review

Lieke M. Reuchlin-Vroklage, MD; Sita Bierma-Zeinstra, PhD; Marc A. Benninga, MD, PhD; Marjolein Y. Berger, MD, PhD

**Background:** Constipation is a common problem in children. Diagnosis is based on clinical features. In case of doubt about the presence of constipation, the existence of fecal retention can be evaluated on plain abdominal radiography.

**Objectives:** To describe and to assess the evidence from observational, controlled studies concerning the association between abdominal radiography and symptoms and signs related to constipation in children.

**Methods:** MEDLINE was searched from inception to April 2004 using a specified search strategy. Studies that fulfilled predefined criteria were assessed for methodological quality. Study characteristics and associations were extracted and the results were summarized according to a best-evidence synthesis.

**Results:** Of the 392 publications identified, 6 studies met the inclusion criteria. Only 2 studies were of high methodological quality. The best-evidence synthesis yielded conflicting evidence for an association between a clinical and a radiological diagnosis of constipation. The likelihood ratio (LR) in 2 high-quality studies was close to 1 (LR, 1.2; 95% confidence interval [CI], 1.0-1.4; and LR, 1.0; 95% CI, 0.5-1.6). Conflicting evidence was found for an association between digital rectal examination and fecal impaction on radiography. Limited evidence was found for an association between a history of hard stool and a finding of rebound tenderness and radiography (LR, 1.2; 95% CI, 1.0-1.4; and LR, 1.1; 95% CI, 1.0-1.2, respectively).

**Conclusions:** The limited amount of data available shows conflicting evidence for an association between clinical symptoms of constipation and fecal loading on abdominal radiographs in children. The recommendation to perform a plain abdominal radiograph in case of doubt of the presence of constipation in a child cannot be supported. Further research of good methodological quality is needed.

Arch Pediatr Adolesc Med. 2005;159:671-678

**CONSTITUTION WITH OR without encopresis is a common problem in children. It affects 3% of preschool-aged children and 1% to 2% of school-aged children.** Peak incidence occurs at the time of toilet training, between the ages of 2 and 4 years. Defecation disorders represent the chief reports in 3% of pediatric outpatient visits and 10% to 25% of pediatric gastroenterology visits. The general well-being of adult patients with chronic constipation is lower than that of a comparable normal population. In children it is well known that encopresis sometimes leads to social withdrawal, low self-esteem, and even depression. Longer duration of symptoms before diagnosis has been associated with poorer long-term outcome.

In more than 90% of the children with constipation and encopresis, no organic or anatomical cause can be found and, therefore, these patients are considered as having a functional defecation disorder. These patients have been later classified by the Pediatric Rome Criteria as having functional constipation, functional fecal retention, or functional nonretentive fecal soiling.

The key features of pediatric constipation are a defecation frequency fewer than 3 times per week in combination with the involuntary loss of stool in the underwear. Other important clinical parameters are stool consistency, large-caliber stools, pain at defecation, and abdominal discomfort. However, there are no well-designed studies that determine which aspects of a medical history and physical examination are most important in discriminating between constipation and non-constipation.

In case of doubt about the presence of constipation one would like to evaluate the
The presence of fecal retention or impaction. More than a decade ago Barr et al. introduced a score to appraise fecal retention on a single radiograph of the abdomen. Since then similar scoring systems have been developed to assess fecal loading on an abdominal radiograph. However, interobserver and intraobserver agreement on the existence of constipation on abdominal radiography varies between the scoring systems used.

To evaluate the additional diagnostic value of abdominal radiography in the diagnosis of constipation in children, we performed a systematic literature review. The objective was to describe and to assess the evidence from observational, controlled studies concerning the association between abdominal radiography and symptoms and signs related to constipation in children.

### Table 1. The QUADAS Instrument*

<table>
<thead>
<tr>
<th>QUADAS Instrument (No. of Studies That Fulfilled This Item)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Was the spectrum of patients representative of the patients who will receive the test in practice? (3)</td>
</tr>
<tr>
<td>2. Were selection criteria clearly described? (3)</td>
</tr>
<tr>
<td>3. Is the reference standard likely to correctly classify the target condition? (5)</td>
</tr>
<tr>
<td>4. Is the period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the 2 tests? (1)</td>
</tr>
<tr>
<td>5. Did the whole sample or a random selection of the sample, receive verification using a reference standard of diagnosis? (5)</td>
</tr>
<tr>
<td>6. Did patients receive the same reference standard regardless of the index test result? (4)</td>
</tr>
<tr>
<td>7. Was the reference standard independent of the index test (ie, the index test did not form part of the reference standard)? (5)</td>
</tr>
<tr>
<td>8. Was the execution of the index test described in sufficient detail to permit replication of the test? (5)</td>
</tr>
<tr>
<td>9. Was the execution of the reference standard described in sufficient detail to permit its replication? (5)</td>
</tr>
<tr>
<td>10. Were the index test results interpreted without knowledge of the results of the reference standard? (3)</td>
</tr>
<tr>
<td>11. Were the reference standard results interpreted without knowledge of the results of the index test? (4)</td>
</tr>
<tr>
<td>12. Were the same clinical data available when test results were interpreted as would be available when the test is used in practice? (2)</td>
</tr>
<tr>
<td>13. Were intermediate test results reported? (0)</td>
</tr>
<tr>
<td>14. Were withdrawals from the study explained? (3)</td>
</tr>
</tbody>
</table>

*Adapted from Whiting et al.13

search strategy for identifying diagnostic studies, as described by Devillé and Buntinx.12 Additional strategies for identifying trials included searching the reference lists of review articles and included studies. Experts in the field were asked to identify further published and unpublished primary studies.

### STUDY SELECTION

Only controlled, observational studies investigating the relationship between fecal loading on plain abdominal radiography and symptoms and signs related to constipation in otherwise healthy children aged from 1 to 18 years were eligible for inclusion in the review. The reference tests considered included fecal loading on plain abdominal radiography according to a predefined scoring system, as well as a clinical diagnosis of constipation according to the presence or absence of predefined symptoms and signs. Results of the study had to allow for the extraction of information on diagnostic value (sensitivity, specificity, likelihood ratio [LR], or accuracy). No language restriction was applied.

Two reviewers (L.M.R.-V. and M.Y.B.) independently screened titles and abstracts of studies identified by the literature search for eligibility. All potentially relevant studies were retrieved as full papers and then again independently reviewed by 2 reviewers (L.M.R.-V. and M.Y.B.). Decisions regarding the inclusion of studies were made independently and any disagreements were resolved through consensus whenever possible, or by arbitration of a third reviewer (S.B.-Z.).

### METHODOLOGICAL QUALITY

Two reviewers (L.M.R.-V. and M.Y.B. or S.B.-Z.) independently assessed the methodological quality of the included studies using the QUADAS instrument (Table 1). All selected methodological criteria were scored as yes (1), no (0), or do not know (0). The overall methodological quality of a study was computed by counting the number of positive scores, with equal weights applied on all items. Scores could range from 0 to 14. We arbitrarily regarded trials with methodological quality scores of 9 or higher (≥60%) as being of high quality. In case of a disagreement between the 2 reviewers, consensus was used to resolve disagreement. When consensus could not be reached, a third reviewer made the final decision (M.Y.B. or S.B.-Z.).

### DATA EXTRACTION

Two reviewers (L.M.R.-V. and M.Y.B.) performed independently a structured extraction of data from the original reports. Disagreements were resolved by consensus. Extracted information included (if available) demographic data, a definition of the participants with and without the disease as determined by the reference test, detailed description of the reference test and index test, reproducibility of the index test, and outcome measures.

### DATA ANALYSIS

The interassessor agreement on the methodological quality was calculated using κ scores. The κ values range from −1, indicating perfect disagreement, to +1, indicating perfect agreement. A κ value greater than 0.7 indicates a high level of agreement between assessors; a κ value between 0.5 and 0.7, a moderate level of agreement; and a κ value less than 0.5, a poor level of agreement. Wherever possible we calculated sensitivities, specificities, and LR with a 95% confidence interval (CI).
In case of clinical heterogeneity (patient population and/or definition of reference and index test are not considered to be sufficiently similar), the results were not pooled but a best-evidence synthesis was used to summarize these data. The level of evidence was then ranked. Strong evidence exists when multiple high-quality studies show consistent findings, moderate evidence exists when one high-quality study and multiple low-quality studies show consistent findings, and limited evidence exists when one high-quality study is found or multiple low-quality studies show consistent findings. Findings are considered consistent when 75% or more of the studies reported the same result. Only statistically significant associations are considered as associated in this synthesis. A statistically significant association exists when the 95% CI of the LR does not include 1.

RESULTS

LITERATURE SELECTION

A total of 392 publications were identified by our search strategy. Of these, 12 studies met our initial inclusion criteria.21,15-23 Seven studies were excluded for the following reasons: in the study of Blethyn et al10 only children with urinary tract infections were included, van der Plas et al24 did not include a control group, in the study of Bewley et al18 no data on diagnostic value were presented, and in 4 studies the symptoms of constipation were not related to the outcome of a plain abdominal radiography.15,19,20,23 No additional studies could be retrieved from the reference lists of relevant articles. Therefore, a total of 6 studies were included in this review.9,11,16,17,21 Full details of the included studies are presented in Table 2 and Table 3.

We included 3 case series,16,17,21 2 case-control studies9,11 and 1 study of a retrospective reexamination of abdominal radiographs.10 In 1 case-series17 the association between the radiographic diagnosis of constipation and the clinical diagnosis of constipation was evaluated in a selection of all included children. In our analysis this study was further considered as a case-control study. All studies were hospital based.

In only 1 study was the purpose to determine whether clinical variables accurately identify children with radiographically proven constipation (Table 3).16 In the 6 studies included, 3 different scoring systems for assessing fecal impaction on abdominal radiography were used. Three studies used the Barr-score,9,17,21 2 studies used a revision of the Barr-score16,19 and Leech et al11 rated according to their own scoring system.

In 3 studies a clinical diagnosis of constipation was considered as the reference test.9,11,17 Only 2 studies gave a good description of the diagnosis of constipation9,17 (Table 3). In the study of Benninga et al17 patients with soiling or encopresis were eligible when they had at least one other symptom of constipation, whereas in the study of Barr et al17 children with soiling or encopresis were excluded. In the 2 case series the abdominal radiograph was considered as the reference test16,21; 1 of these 2 studies retrospectively retrieved its information on symptoms and signs of constipation from medical record forms.21 In the latter study only children with encopresis were included.

Table 2. Purposes of the Included Studies

<table>
<thead>
<tr>
<th>Source</th>
<th>Study Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beckmann et al16</td>
<td>To determine whether clinical variables accurately identify children with radiographically proven constipation</td>
</tr>
<tr>
<td>Leech et al11</td>
<td>To assess the reliability of scoring fecal loading on plain abdominal radiographs in children with intractable constipation</td>
</tr>
<tr>
<td>Blethyn et al19</td>
<td>To determine whether the degree of fecal loading could be reliable and reproducible when assessed by different observers using a single abdominal radiograph</td>
</tr>
<tr>
<td>Rockney et al27</td>
<td>To determine whether (1) fecal retention in encopretic children can be assessed objectively using a plain abdominal radiograph; and (2) radiological evidence of fecal retention is associated with clinical findings on presentation in encopretic children</td>
</tr>
<tr>
<td>Benninga et al17</td>
<td>To assess the presence or absence of fecal retention in children using CTT; to compare findings on CTT with a plain abdominal radiograph scored according to Barr</td>
</tr>
<tr>
<td>Barr et al9</td>
<td>To develop a method to assess severity of stool retention using plain abdominal radiographs</td>
</tr>
</tbody>
</table>

Abbreviation: CTT, colonic transit time.

QUALITY ASSESSMENT

The reviewers scored 84 items and agreed on 65 items (77.4%, χ² = 0.54) (Table 1). The 19 disagreements were resolved in a single consensus meeting. Disagreement was largest at item 4 (50%): “Is the period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the 2 tests?” The total scores ranged from 5 to 11. Items 4, 12, and 13 were the least fulfilled criteria (1, 2, and 0 times, respectively). The mean methodological quality score was 8.3.

DIAGNOSTIC VALUE

Beckmann et al16 performed a multivariate discriminant analysis. In a logistic regression model, clinical symptoms of constipation were introduced as the independent variables and radiographic constipation as the dependent variable. The model defined a clinical diagnosis of constipation. This clinical diagnosis existed 1.2 times as often in children with a radiographic diagnosis of constipation as in children without a radiographic diagnosis of constipation (LR, 1.2; 95% CI, 1.0-1.4) (Table 4).

Four studies examined whether abdominal radiography was able to discriminate between children with clinical constipation and children without clinical constipation.9,11,17 Only Leech et al11 found a statistically significant discriminative value (LR, 3.0; 95% CI, 1.6-4.3) (Table 4). Blethyn et al19 presented accuracy as a measure for the ability of abdominal radiography to discriminate between clinically constipated and nonconstipated children (accuracy, 80%; 95% CI, 50%-100%).
Table 3. Study Characteristics

<table>
<thead>
<tr>
<th>Source</th>
<th>Participants</th>
<th>Reference Test</th>
<th>Index Test</th>
<th>Reliability of the Score of the Abdominal Radiograph</th>
<th>Quality Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beckmann et al 16</td>
<td>Children presenting with abdominal pain to the ED of Children's Hospital of Wisconsin, Milwaukee, between January 1997 and January 1998 who underwent abdominal radiography based on customary practices. Children with previous abdominal surgery, known abdominal pathology, menarche, or sickle cell disease were excluded. In total 410 children were eligible, but all the data of only 251 children were available. N = 251; mean age, 7 y (age range, 4.4-11 y; 48% male)</td>
<td>Fecal loading on abdominal radiograph according to Blethyn et al. 10. Grade 0: Feces located in the rectum and cecum only. Grade 1: Feces in the rectum, cecum and discontinuous elsewhere. Grade 2: Feces in the rectum and cecum with continuous feces affecting all segments but allowing for gas. Grade 3: Continuous feces with dilated colon and rectal impaction. Radiographically proven constipation defined as grade 1-3. Cases: Children with a clinical diagnosis of constipation (not further specified); N = 33; 70% male; age range, 3 mo to 13.5 y Controls: Children who underwent IVP for suspected renal tract disorder. N = 67; 63% male; age range, 1 mo to 14 y It remains unclear as to how many children younger than age 1 y are included. Clinical diagnosis of constipation based on bowel frequency.</td>
<td>Structured data sheet assessing: Symptoms: History of gastrointestinal tract problems, duration of abdominal pain, stooling habits, straining with defecation, fecal consistency, medication. Signs: Physical examination of the abdomen: rebound tenderness, rigidity, guarding, tympanic, distension, tenderness in different regions of the abdomen, bowel sounds, and rectal examination. Demographic information. Rationale for radiography was the initial diagnosis. Abdominal radiography divided in 3 segments, each segment given a score from 0 to 5, giving a total score of 0-15. 0 indicates no feces visible; 1, scanty feces visible; 2, mild fecal loading; 3, moderate fecal loading; 4, severe fecal loading; and 5, severe fecal loading with bowel dilatation. Total score 8-15 indicates significant constipation. Abdominal radiography graded according to a revision of the Barr-score. Grade 0: Feces in rectum and cecum only. Grade 1: Feces in rectum and cecum and discontinuous elsewhere. Grade 2: Feces in rectum and cecum and continuous (allowing for gas), affecting all segments. Grade 3: Feces in rectum and cecum and continuous elsewhere with dilated colon and impacted rectum.</td>
<td>Not mentioned</td>
<td>11</td>
</tr>
<tr>
<td>Leech et al 17</td>
<td>Children referred to the gastroenterology clinic at the John Radcliffe Hospital, Oxford, UK, between September 1994 and September 1995. Underlying neurological and spinal abnormalities were excluded.</td>
<td>Abdominal radiography scored according to Barr: Total score: 0-25. Score of $\geq$10 indicates fecal retention.</td>
<td>Intraobserver variation by Wilcoxon matched-pairs signed ranks test: $P$ = .12-0.69. Interobserver variation by Friedman 2-way analysis of variance $\chi^2$ = 44.205, $df$ = 2, $P$&lt;.05.</td>
<td>5</td>
<td></td>
</tr>
<tr>
<td>Blethyn et al 10</td>
<td>Children who underwent abdominal radiography for a variety of clinical indications including constipation, urinary tract infection, and hematuria. N = 20; age range 5-14 y In 15 children information on bowel frequency was available.</td>
<td>Abdominal radiography scored according to Barr: Total score: 0-25. Score of $\geq$10 indicates fecal retention.</td>
<td>Intraobserver variation between 2 radiologists with pediatric experience (agreement 18 of 20): $\kappa$ = 0.95, $P$&lt;.001. Interobserver variation between 4 less experienced radiologists with 1 experienced radiologist (agreement 60 of 80): $\kappa$ = 0.8434, $P$&lt;.001.</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>Rockney et al 13</td>
<td>Children referred to 2 incontinence clinics, 1 tertiary care hospital, and 1 community hospital for evaluation and management of fecal or urinary incontinence for whom an abdominal radiograph was available. Physical disorders such as aganglionic megacolon were excluded. N = 60; age range 4-18 y; 72% male. 92% of the children could be diagnosed as having encopresis according to DSM III.</td>
<td>Abdominal radiography scored according to Barr: Total score: 0-25. Score of $\geq$10 indicates fecal retention.</td>
<td>Interobserver reliability: Overall agreement of the systematic assessment of fecal retention among 3 radiologists is 87.4% ($\kappa$ = 0.83, $P$&lt;.0001). Intraobserver reliability for the systematic assessment: ($\kappa$ range, 0.52-0.63).</td>
<td>8</td>
<td></td>
</tr>
</tbody>
</table>

(continued)
### Table 3. Study Characteristics (cont)

<table>
<thead>
<tr>
<th>Source</th>
<th>Participants</th>
<th>Reference Test</th>
<th>Index Test</th>
<th>Reliability of the Score of the Abdominal Radiograph</th>
<th>Quality Score (n = 14)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benninga et al17</td>
<td>Otherwise healthy children, referred to the Department of Pediatric Gastroenterology of the Academic Medical Centre, Amsterdam, the Netherlands, for complaints of infrequent defecation, soiling, encopresis, or recurrent abdominal pain, between May 1991 and April 1993. Children with Hirschsprung disease, spinal and/or anal anomalies, prior surgery of the colon, metabolic diseases, mental retardation, or using drugs other than laxatives, were excluded. Total included N = 211; age range, 5-17 y. 101 Underwent abdominal radiography.</td>
<td>Cases of PC: At least 2 of the following 4 criteria: Stool frequency &lt;3 times per week; ≥2 soiling/encopresis episodes per week; periodic passage of very large amounts of stools once every 7-30 d; a palpable abdominal or rectal mass (N = 57). Controls: Solitary encopresis and/or soiling: without any of the other criteria of PC (N = 30). Recurrent abdominal pain: Severe enough to interfere with day-to-day activities over at least a 3-mo period without any of the other symptoms of PC (N = 14).</td>
<td>Abdominal radiography scored according to Barr: Total score: 0-25; score of ≥10 indicates fecal retention.</td>
<td>Interobserver agreement of 2 radiologists for 2 separate radiographs scoring different colon segments: κ = 0.28 in the first radiograph, and κ = 0.60 in the second radiograph. Interobserver agreement on the existence of constipation as measured by a Barr-score of 10 between 2 radiographs: κ = 0.22, and κ = 0.25 respectively, for the 2 radiologists.</td>
<td></td>
</tr>
<tr>
<td>Barr et al8</td>
<td>All children presenting to an ambulatory pediatric clinic in an 11-mo period with reports of constipation or recurrent abdominal pain were included when (1) pretherapy and posttherapy films were available and (2) sufficient follow-up information was available to judge 'success' or 'failure' of therapy at the time of the follow-up radiological study. Included were 44 abdominal radiographs of children suspected of fecal loading and 14 abdominal radiographs of children suspected to have no fecal loading.</td>
<td>Cases: Children presenting with constipation or recurrent abdominal pain in whom symptomatic stool retention was diagnosed, based on evidence of “pellet” stools, straining, having a bowel movement no more often than every 3 d, blood streaking on stools, very large stools, history of soiling, positive rectal examination or colonic stool palpated on abdominal examination. Patients with a present history of soiling were excluded; N = 30; age not specified. Controls: Children who had abdominal radiography for lead ingestion and who did not present with either abdominal pain or constipation and who had blood lead levels &lt;50 µg/dL (2.41 µmol/L); N = 12; age range, 3-7 y. Success: Absence of symptoms with which the patient presented following therapy; N = 19. Failure: Persistence of symptoms with which the patient presented following therapy; N = 14. Cases and failures were those that reference test–positive results. Controls and successes were those with reference test–negative results.</td>
<td>Abdominal radiography scored according to Barr: Score of quantity of feces in ascending colon, 0-2; transverse colon, 0-5; descending colon, 0-5; and rectum: 0-5. Score quality of stool: Rocky feces proximally, 0-3; granular feces distally, 0-5. Total score: 0-25; a score of ≥10 indicated fecal retention.</td>
<td>Interobserver reliability, ≥0.80. Intraobserver reliability, ≥0.85.</td>
<td></td>
</tr>
</tbody>
</table>

**Abbreviations:** DSM-III, Diagnostic and Statistical Manual of Mental Disorders, Third Edition; ED, emergency department; IVP, intravenous pyelogram; PC, pediatric constipation.
Diagnostic value was related to the methodological quality of the studies. Low-quality studies\textsuperscript{9,11} reported better diagnostic values compared with high-quality studies\textsuperscript{16,17} (Figure). In particular the specificity was poor in the high-quality studies (Table 4).

The diagnostic value of individual symptoms was evaluated in 2 studies\textsuperscript{16,21}. Results are presented in Table 5. Beckmann et al\textsuperscript{16} defined a positive rectal examination as stool present in the rectal vault. They found a significant LR for a finding of stool at rectal examination and radiographic constipation (limited evidence).\textsuperscript{16} One high-quality study\textsuperscript{17} and 1 low-quality study\textsuperscript{9} did not find a statistically significant association between a clinical diagnosis of constipation and a radiographic diagnosis of constipation (50% of the findings). One high-quality study\textsuperscript{17} and 1 low-quality study\textsuperscript{9} did not find a statistically significant association (50% of the findings). These findings are inconsistent (conflicting evidence). One high-quality study found a significant association between a finding of stool at rectal examination and radiographic finding of constipation.\textsuperscript{16} This association was not statistically significant in another low-quality study (conflicting evidence).\textsuperscript{21} One high-quality study found a significant association between rebound tenderness and radiographic constipation (limited evidence).\textsuperscript{16}

### INTEROBSERVER AND INTRAOBSERVER RELIABILITY

Only in the study of Benninga et al\textsuperscript{17} did interobserver reliability of the rating of the radiographs range from poor to moderate ($\kappa=0.28-0.60$). All other studies reported moderate to excellent interobserver reliability ($\kappa$ range, 0.63\textsuperscript{21-0.95}\textsuperscript{10}) (Table 3). Intraobserver reliability was evaluated in 3 studies\textsuperscript{9,11,21} and ranged from moderate ($\kappa=0.52$) to excellent ($\kappa=0.85$) (Table 3).

### BEST-EVIDENCE SYNTHESIS

This systematic review of the literature was performed to investigate the possible association between clinical symptoms and radiographic features of childhood constipation. We identified a total of 6 studies, that were eligible for inclusion in this study. Only 2 of these studies were of good methodological quality.\textsuperscript{16,17}

We found conflicting evidence for an association between a clinical and a radiographic diagnosis of constipation. Methodological quality of the study was related to the diagnostic value reported by the studies. High-quality studies (score $\geq 9$) found LRs close to 1 in contrast to low-quality studies that reported high LRs (Figure).
The results of rectal examination were not consistently related to findings on fecal retention at abdominal radiography. Likelihood ratios of a finding of stool at rectal examination were close to 1, indicating that stool at rectal examination occurred almost as often in children with fecal loading on radiography as in children without fecal loading (Table 5). The time between rectal examination and abdominal radiography might have influenced the accuracy of the abdominal radiography. In the high-quality study abdominal radiography was performed almost directly after rectal examination.21 In the low-quality study this period was not specified.21

The interreliability and intrarelability of scoring systems for fecal loading on abdominal radiography varied among studies from poor to excellent. Experienced radiologists had more reliable scores than less experienced colleagues. Benninga et al17 investigated the reproducibility of the Barr-score on 2 consecutive radiographs 3 days apart, without administering a laxative. The interobserver reliability between the 2 radiographs on the existence of constipation as measured by a Barr-score of 10 was poor. In their study the radiographic diagnosis of constipation in a child was not reproducible.

**CLINICAL IMPLICATIONS**

Constipation can be diagnosed by a detailed medical history and a thorough physical examination including a digital rectal examination. Symptoms of constipation like stool frequency, encopretic episodes, and the passage of large amounts of stools are reliably recalled by the child or the child’s parents.29 Nevertheless a diagnosis might be doubtful in case not all key symptoms of constipation are present or when a rectal examination is infeasible. The clinical practice guideline of the North American Society for Pediatric Gastroenterology and Nutrition for diagnosing constipation in children recommends taking a plain abdominal radiograph in case of doubt about the presence of constipation, in case the child refuses a rectal examination, or if rectal examination is considered traumatic.8 The results of our review do not support this recommendation. We found conflicting evidence for an association between a clinical diagnosis of constipation and a radiographic diagnosis. Furthermore, high-quality studies found LRs close to 1 indicating that a radiographic diagnosis of constipation occurs almost as often in clinically constipated as in clinically nonconstipated children.

Although there was wide heterogeneity between the 6 studies, all studies were homogeneous in their hospital-based setting. Therefore, the results of our review cannot be generalized to general practice.

**LIMITATIONS OF THIS REVIEW**

Although the literature search was extensive and had no language restriction, it might have a risk of publication bias. We found 6 studies of which only 2 were of good methodological quality. Therefore, one missed high-quality study may easily alter the results of our best-evidence synthesis. However, because none of the consulted experts in the field could add any missed or ongoing study, it is unlikely that such a study exists.

In this review only statistically significant LRs were considered of diagnostic value. This criterion may misclassify the results of studies with a small sample size (introduction of type II error due to the low statistical power of small studies). The conflicting evidence we found for the association between a clinical and a radiographic diagnosis of constipation was not influenced by the sample size of the studies. In contrast, small sample size may be of influence in our finding of conflicting evidence for an association between stool on rectal examination and radiographically confirmed constipation. In a study of 251 children Beckmann et al26 reported a statistically significant LR of 1.6 (95% CI, 1.2-2.0); Rockney et al21 studied 60 children and reported an LR of 1.5 but did not reach statistical significance (95% CI, 0.8-2.3). Although pooling of data would have been a solution to overcome the problem of small sample size, we refrained from pooling because of the substantial differences between studies.

**FUTURE RESEARCH**

Most clinicians consider a clinical diagnosis of constipation as the gold standard. However, there are no well-designed studies that determine which aspects of a medical history and physical examination are most important in discriminating between constipation and nonconstipation.8 de Lorijn et al25 found a good relation between clinical symptoms and colonic transit time. However, 50% of their patients who fulfilled the clinical criteria for constipation had colonic transit times within the normal range. Therefore, one might argue that a clinical diagnosis of constipation as well as fecal loading on radiographs are both substitutes of an adequate reference standard for constipation in children. In case of lack of an adequate reference standard, follow-up studies (preferably randomized) to quantify the effect of a diagnostic test on patient outcome are needed. Evaluating a test on patient outcome involves the evaluation of the diagnostic tests (clinical diagnosis, abdominal radiography, and colonic transit time) plus current administered therapies (laxatives and behavioral interventions) combined.26 Given the LRs of almost 1 in the high-quality studies, one should discuss whether the diagnostic strategy evaluated in such a study should incorporate plain abdominal radiography.

**CONCLUSIONS**

In the limited data available we found conflicting evidence for a diagnostic association between clinical symptoms of constipation and fecal loading in abdominal radiographs in children. The recommendation to perform a plain abdominal radiograph in case of doubt of the presence of constipation in a child cannot be supported by this systematic review. Further well-powered research of good methodological quality is needed to find the best diagnostic strategy in children suspected of having constipation.
Accepted for Publication: February 13, 2005.
Correspondence: M. Y. Berger, Department of General Practice, Erasmus Medical Centre, K FF37, PO Box 1738, 3000DR Rotterdam, the Netherlands (m.berger@erasmusmc.nl).

REFERENCES