Published in final edited form as:

J Pediatr Gastroenterol Nutr. 2010 August ; 51(2): 155–159. doi:10.1097/MPG.0b013e3181cb4309.

LACK OF UTILITY OF ABDOMINAL X-RAYS IN THE EVALUATION OF CHILDREN WITH CONSTIPATION: COMPARISON OF DIFFERENT SCORING METHODS

L Pensabene^{*}, C Buonomo, L Fishman, D Chitkara, and S Nurko

Center for Motility and Functional Gastrointestinal Disorders, Children's Hospital, Boston, MA, United States.

Abstract

Abdominal X-rays are used diagnostically in the evaluation of children with constipation. However, their clinical utility has not been established.

Aim—To assess the accuracy of different methods in identifying children with functional constipation (FC) or non-retentive fecal incontinence (NRFI).

Methods—Retrospective review of abdominal x-rays in which colonic transit (CT), Barr, Leech and fecal loading (FL) scores were blindly measured by blinded pediatric gastroenterologists and a radiologist. Children were classified *a priori* as FC or NRFI.

Results—160 patients (125 FC, 35 NRFI) were studied. There were significant differences ($p < 0.05$) when comparing those with FC and those with NRFI: CT: 51 ± 18 vs 40 ± 21 hrs; Barr: 14 ± 5 vs 11 ± 4 ; Leech: 10 ± 2 vs 8 ± 2 ; FL: 2 ± 0.5 vs 1.7 ± 0.4 . Over 20% of FC, had normal Barr and Leech scores, while $> 50\%$ of NRFI had abnormal scores. CT discriminated better between FC and NRFI. There was a significant correlation ($p < 0.05$) between CT and Barr (0.45), Leech (0.41) and FL scores (0.36), and between Barr and Leech scores ($r = 0.94$). There was good intraobserver correlation between Barr, Leech and FL scores but poor interobserver reproducibility.

Conclusions—While significant differences in overall FC and NRFI scores exist, the discriminative value is low for all scores. There is poor interobserver reproducibility of the Barr, Leech and FL scores. These findings confirm the limited value of the plain abdominal X-ray in the evaluation of children with constipation.

INTRODUCTION

One of the most common functional gastrointestinal disorders (FGIDs) in childhood is chronic constipation (1–4). The Rome III FGIDs classification (5–7), and the Paris Consensus on Childhood Constipation Terminology (PACCT) (8), state that the identification and classification of functional constipation should be based on symptoms and clinical history. However, a plain abdominal radiograph in which fecal loading is

Correspondence: Samuel Nurko MD, Children's Hospital, 300 Longwood Ave, Boston Ma 02115, Phone: 617 355 6055, Fax: 617 730 0043, Samuel.nurko@childrens.harvard.edu .

^{*}Presently at: Department of Pediatrics, University Magna Graecia, Catanzaro, Italy

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

determined, is still frequently used in the evaluation of children with constipation, either to support the diagnosis, or to suggest the type or response to treatment (9). Plain visual inspection is highly subjective, and its clinical utility has not been established, as suggested in the evidence based guidelines published by the North American Society of Pediatric Gastroenterology, Hepatology and Nutrition (NASPGHAN) (9). Limited studies and a recently published systematic review have indicated that the plain abdominal radiograph to assess fecal loading has high inter- and intra-observer variability and poor diagnostic accuracy (9–15). Therefore to make the evaluation of fecal loading in abdominal radiograph less subjective different scoring systems have been developed (11,16). However there is limited information on the use and utility of these different scoring methods during daily clinical work (9,13,14,17–20).

The aim of this study was to assess and compare the accuracy of different radiographic scoring methods for the identification of children with functional constipation.

MATERIAL AND METHODS

Study population

All children with defecation complaints who were referred to the Pediatric Gastrointestinal Motility Unit at Children's Hospital in Boston from 2000–2004, that had undergone a colonic transit study as part of their clinical evaluation were eligible for this study. Patients with organic disease, known causes of constipation (Hirschsprung's disease, hypothyroidism, myelomeningocele metabolic or renal diseases) or fecal impaction as defined by the PACTT criteria (8) were excluded. The study was approved by the Institutional Review Board.

Clinical diagnosis

Children were classified *a priori*, without any knowledge of the colonic transit time, by two blinded investigators, as having either functional constipation (FC) or non-retentive fecal incontinence (NRFI) using the PACTT (8), and Rome III criteria (6,7). The diagnosis of *Functional constipation (FC)* was based on the following criteria:

The occurrence of two or more of the following characteristics for at least 8 weeks:

- Frequency of bowel movements less than three per week
- At least one episode of fecal incontinence per week
- Large stools in the rectum or palpable on abdominal examination
- Passing of stools so large that they may obstruct the toilet
- Display of retentive posturing and with-holding behaviors
- Painful defecation

Patients with FC were further divided in two groups according to the presence or absence of fecal incontinence.

The diagnosis of *Non-retentive fecal incontinence (NRFI)* was based on the following criteria:

- The passage of stools in an inappropriate place for at least 8 weeks, occurring in children with a mental age of 4 years and older, with no evidence of constipation based on history and/or examination

Abdominal radiography and scoring methods

De-identified abdominal x-rays performed for colonic transit (CT) measurement were retrospectively reviewed by 3 blinded pediatric gastroenterologists (SN, LF, DC) and a pediatric radiologist (CB) and scores for colonic transit (CT), Barr, Leech and subjective amount of fecal loading (FL) were measured. The x-rays were shown to each investigator in a random fashion. Prior to scoring any films, the methods for the Barr and Leech scores were reviewed and discussed until both investigators agreed on the characteristics of each scoring element (see Scoring). Then the Barr and Leech scores were independently calculated by one pediatric gastroenterologist (SN) and the radiologist (CB). Three pediatric gastroenterologists (SN, LF,DC) independently scored the amount of fecal loading present in each radiograph. No attempt to standardize this subjective score was undertaken. The colonic transit time was assessed by a single independent scorer (LP) in duplicate to check for accuracy as it was assumed that counting of radiopaque markers would not lead to intra- or interobserver variability.

Colonic transit—All laxatives were stopped at least 48 hours before the colonic transit was done. Therefore the colonic transit was performed off medications. All patients had a physical examination before the colonic transit was requested, to ensure there was no fecal impaction (as defined by the PACTT criteria) (8). Therefore no patient with a fecal impaction was included.

The colonic transit study was done following a modification of the technique reported by Metcalf et al (21): patients ingested one capsule containing 24 radio-opaque markers (Sitzmarks. Konsyl Pharmaceuticals, Inc. Fort Worth, Texas) at a similar time each morning for three consecutive days; on the fourth day a plain abdominal radiograph was obtained at the same time of the day (21). The abdomen was then divided into 3 segments as outlined by Arhan (22). Markers were counted in each segment. Total transit time was calculated using the formula described by Metcalf et al., adjusting for the fact that 24 markers were ingested each day instead of 20 as in the original formula. This adjustment allows the calculation of transit times (in hours) by simply counting the markers present in the abdominal radiograph taken on the fourth day. The reference values used were those of Arhan in healthy children. The presence of > 62 markers was considered abnormal (22).

Leech score—This score divides colon into three segments: Right colon, left colon and rectosigmoid segment. Each segment is provided with a score from 0 to 5; 0=no feces visible, 1=scanty feces, 2+ mild fecal loading; 3=moderate fecal loading; 4=severe fecal loading; 5=severe fecal loading with bowel dilatation (11). Scores range from 0 to 15. A score of 9 or more was considered abnormal (11).

Barr score—This score quantifies the amount of feces in 4 bowel segments (ascending colon, transverse colon, descending colon and rectum), and also the consistency of the feces (10). A score of > 10 was considered abnormal.

Fecal loading—This was a subjective score in which the investigators classified the amount of stool in the radiograph as mild (1 point), moderate (2 points) or severe (3 points), in accordance with their clinical practice. A score of > 2 was considered abnormal.

Statistical analyses

All results are expressed as mean $\pm$ SE. Parametric and non-parametric tests were used to compare the different patient characteristics and the scores between the two groups. The reported scores represent the averages obtained of the individual scores. A subgroup comparison was performed in the group with FC according to the presence or absence of

fecal incontinence. Sensitivity and specificity were calculated; kappa coefficients were calculated as indicators of inter- and intra-observer agreement for categorical outcome variables. Receiver Operator Characteristic curves (ROC curves) were created, and the area under the curve (AUC) and 95% confidence intervals calculated. From the AUC the diagnostic accuracy was graded as low (0.5–0.7), moderate (0.7 to 0.9) or high (> 0.9) (23)

RESULTS

There were 160 consecutive patients with functional defecation disorders who underwent abdominal x-rays for colonic transit (CT). Within this cohort, a total of 125 children fulfilled the clinical criteria for FC and 35 were diagnosed as having NRFI. Table 1 shows baseline characteristics of the studied patients.

Comparison of scores between children with FC or NRFI

There were significant differences ($p < 0.05$) among different scores when comparing children with FC to those with NRFI (figure 1). Distribution of normal and abnormal scores in the two patient groups is shown in table 2. While Barr and Leech scores were abnormal more often in children with FC, it should be noted that >20% of FC had normal scores, and > 50% of NRFI had abnormal scores.

Comparison of scores in children with FC between those with or without fecal incontinence

Scores in patients with functional constipation with and without fecal incontinence are shown in figure 2. Patients with constipation and fecal incontinence had significantly higher scores as compared with those with constipation alone ($p < 0.01$).

Characteristics of the scores

Correlation between the different scoring methods is shown in table 3. There was a good correlation between the different methods. There was also a good intraobserver correlation between the Barr, Leech and FL scores. However, there was a poor interobserver reproducibility of the Barr (kappa 0.3), Leech (kappa 0.25) and Fecal loading scores (kappa 0.3). Figure 3 shows ROC curves obtained by the different methods. The AUC were: Barr score 0.672 (0.575 – 0.770 95% confidence interval); Leech score 0.661 (0.564 to 0.759), Fecal loading 0.626 (0.522 to 0.730) and for colonic transit 0.643 (0.535 to 0.751)

DISCUSSION

Functional constipation and NRFI are symptom based diagnosis (6,7,9). However, in clinical practice the existence of fecal retention is frequently evaluated on plain abdominal radiographs and used in determination of diagnosis and therapy (15). We found that even though there were significant differences in CTT and all visual scores when comparing patients with FC to those with NRFI, all scoring methods had low sensitivity, low inter-observer reproducibility and a low discriminative value to differentiate patients with FC from those with NRFI. In our study the subjective assessment of FL by experienced gastroenterologists had the worst performance in respect to its ability to differentiate between groups. This is an important previously unpublished observation, as in clinical practice this subjective non-standardized assessment is the method most commonly used to evaluate children with defecation abnormalities. Our findings suggest that if an abdominal x-ray is used in the evaluation of children with defecation problems, either the Barr or the Leech score will provide more reliable information, although all scores have a low diagnostic accuracy (23).

There was good correlation between the different scoring methods and a good intraobserver correlation, but a low inter-observer reproducibility. While experienced radiologists score more reliably than less experienced colleagues (24), high inter- and intra-observer variability and poor diagnostic accuracy of assessing fecal loading on an abdominal radiograph have also been previously reported (10–14). In the present study, all observers were skilled and experienced in reading plain abdominal radiographs, yet showed the same poor inter-observer reproducibility. However, the fact that there is a good intra-observer correlation suggests that there may be some value in the longitudinal follow up and evaluation of the patient, as long as the same person is interpreting the x-ray every time.

The poor diagnostic accuracy of scoring a single abdominal radiograph may be due to the fact that it represents a single observation in time. The amount of fecal loading is subject to daily variation, depending on food intake and timing of the last defecation. Another possibility for the low sensitivity of the different scoring methods may be that the Rome III and PACCT criteria, are too broad for describing fecal disorders in children. Future research performed in unselected populations would be useful in addressing this question.

Since an abdominal radiograph has low sensitivity and requires the exposure of a child to a small dose of ionizing radiation, we do not advocate its use in the routine evaluation of children with constipation and other defecation abnormalities. An abdominal radiograph may be necessary in obese patients (9), if the child refuses a rectal examination (9), if there is suspicion of sexual abuse (25) or other psychological factors that would make rectal examination inappropriate or too traumatic (26). However the findings of the radiograph need to be interpreted with caution.

There are some limitations to our study. We do not have concurrent physical examination data to correlate with the findings of the x-ray. However a recently published systematic review (15) reported that most studies that have tried to assess if there is an association between clinical symptoms and fecal loading on radiographs in children have showed a lack of correlation. Another limitation is that we do not know the clinical impact that the interpretation of the x-ray had on the patient. We also do not have follow up information to know if our clinical diagnosis remained the same or changed over time. Prospective studies to validate the new Rome III criteria, and to correlate clinical findings with finding on x-ray, as well as outcome studies based on x-ray findings are needed.

In conclusion, the discriminative value to detect functional constipation was low for all scoring methods and there was poor interobserver reproducibility of the Barr, Leech and FL scores. These findings confirm the limited value of the plain abdominal X-ray in the evaluation of children with constipation: use of radiographs is not a substitute for careful history taking and physical examination that remain the cornerstone in the diagnostic work-up of children with functional defecation disorders.

Acknowledgments

This study was supported In part by NIH grant K24DK82792-1 (Dr Nurko)

REFERENCES

1. Loening-Baucke V. Chronic constipation in children. *Gastroenterology* 1993;105(5):1557–1564. [PubMed: 8224663]
2. Partin JC, Hamill SK, Fischel JE, Partin JS. Painful defecation and fecal soiling in children. *Pediatrics* 1992;89(6 Pt 1):1007–1009. [PubMed: 1594338]
3. Loening-Baucke V. Prevalence, symptoms and outcome of constipation in infants and toddlers. *J Pediatr* 2005;146(3):359–363. [PubMed: 15756220]

4. Boccia G, Manguso F, Coccorullo P, Masi P, Pensabene L, Staiano A. Functional defecation disorders in children: PACCT criteria versus Rome II criteria. *J Pediatr* 2007;151(4):394–398. 8 e1. [PubMed: 17889076]
5. Rasquin-Weber A, Hyman PE, Cucchiara S, Fleisher DR, Hyams JS, Milla PJ, et al. Childhood functional gastrointestinal disorders. *Gut* 1999;45 Suppl 2:II60–II68. [PubMed: 10457047]
6. Hyman PE, Milla PJ, Benninga MA, Davidson GP, Fleisher DF, Taminiau J. Childhood functional gastrointestinal disorders: neonate/toddler. *Gastroenterology* 2006;130(5):1519–1526. [PubMed: 16678565]
7. Rasquin A, Di Lorenzo C, Forbes D, Guiraldes E, Hyams JS, Staiano A, et al. Childhood functional gastrointestinal disorders: child/adolescent. *Gastroenterology* 2006;130(5):1527–1537. [PubMed: 16678566]
8. Benninga M, Candy DC, Catto-Smith AG, Clayden G, Loening-Baucke V, Di Lorenzo C, et al. The Paris Consensus on Childhood Constipation Terminology (PACCT) Group. *J Pediatr Gastroenterol Nutr* 2005;40(3):273–275. [PubMed: 15735478]
9. Baker S, Liptak G, Colletti R, Croffie J, Di Lorenzo C, Ector W, et al. Constipation in infants and children : Evaluation and treatment. A medical position statement of the North American Society for Pediatric Gastroenterology Hepatology and Nutrition. *J Pediatr Gastroenterol Nutr* 2006;43:e1–e13. [PubMed: 16954945]
10. Barr RG, Levine MD, Wilkinson RH, Mulvihill D. Chronic and occult stool retention: a clinical tool for its evaluation in school-aged children. *Clin Pediatr (Phila)* 1979;18(11):674. 6, 7–9, passim. [PubMed: 498690]
11. Leech SC, McHugh K, Sullivan PB. Evaluation of a method of assessing faecal loading on plain abdominal radiographs in children. *Pediatr Radiol* 1999;29(4):255–258. [PubMed: 10199902]
12. Blethyn AJ, Verrier Jones K, Newcombe R, Roberts GM, Jenkins HR. Radiological assessment of constipation. *Arch Dis Child* 1995;73(6):532–533. [PubMed: 8546512]
13. Benninga MA, Buller HA, Staalman CR, Gubler FM, Bossuyt PM, van der Plas RN, et al. Defaecation disorders in children, colonic transit time versus the Barr-score. *Eur J Pediatr* 1995;154(4):277–284. [PubMed: 7607277]
14. Jackson CR, Lee RE, Wylie AB, Adams C, Jaffray B. Diagnostic accuracy of the Barr and Blethyn radiological scoring systems for childhood constipation assessed using colonic transit time as the gold standard. *Pediatr Radiol* 2009;39(7):664–667. [PubMed: 19277628]
15. Reuchlin-Vroklage LM, Bierma-Zeinstra S, Benninga MA, Berger MY. Diagnostic value of abdominal radiography in constipated children: a systematic review. *Arch Pediatr Adolesc Med* 2005;159(7):671–678. [PubMed: 15997002]
16. Blethyn AJ, Jenkins HR, Roberts R, Verrier Jones K. Radiological evidence of constipation in urinary tract infection. *Arch Dis Child* 1995;73(6):534–535. [PubMed: 8546513]
17. Zaslavsky C, da Silveira TR, Maguilnik I. Total and segmental colonic transit time with radio-opaque markers in adolescents with functional constipation. *J Pediatr Gastroenterol Nutr* 1998;27(2):138–142. [PubMed: 9702642]
18. Bouchoucha M, Devroede G, Arhan P, Strom B, Weber J, Cugnenc PH, et al. What is the meaning of colorectal transit time measurement? *Dis Colon Rectum* 1992;35(8):773–782. [PubMed: 1644002]
19. Papadopoulou A, Clayden GS, Booth IW. The clinical value of solid marker transit studies in childhood constipation and soiling. *Eur J Pediatr* 1994;153(8):560–564. [PubMed: 7957401]
20. de Lorijn F, van Wijk MP, Reitsma JB, van Ginkel R, Taminiau JA, Benninga MA. Prognosis of constipation: clinical factors and colonic transit time. *Arch Dis Child* 2004;89(8):723–727. [PubMed: 15269069]
21. Metcalf AM, Phillips SF, Zinsmeister AR, MacCarty RL, Beart RW, Wolff BG. Simplified assessment of segmental colonic transit. *Gastroenterology* 1987;92(1):40–47. [PubMed: 3023168]
22. Arhan P, Devroede G, Jehannin B, Lanza M, Faverdin C, Dornic C, et al. Segmental colonic transit time. *Dis Colon Rectum* 1981;24(8):625–629. [PubMed: 7318630]
23. Collinson P. Of bombers, radiologists, and cardiologists: time to ROC. *Heart* 1998;80:215–217. [PubMed: 9875074]

24. Gutierrez C, Marco A, Nogales A, Tebar R. Total and segmental colonic transit time and anorectal manometry in children with chronic idiopathic constipation. *J Pediatr Gastroenterol Nutr* 2002;35(1):31–38. [PubMed: 12142807]
25. Benninga MA, Buller HA, Heymans HS, Tytgat GN, Taminiu JA. Is encopresis always the result of constipation? *Arch Dis Child* 1994;71(3):186–193. [PubMed: 7979489]
26. Bongers ME, Voskuil WP, van Rijn RR, Benninga MA. The value of the abdominal radiograph in children with functional gastrointestinal disorders. *Eur J Radiol* 2006;59(1):8–13. [PubMed: 16632293]

Figure 1. Comparison among children with FC and those with NRFI. There were significant differences among different scores when comparing children with FC to those with NRFI * $p < 0.05$. FC; Functional constipation; NRFI.- Non retentive fecal incontinence

Figure 2. Comparison of scores in patients with functional constipation with or without fecal incontinence. Patients with constipation and fecal incontinence had significantly higher scores as compared with those with constipation alone. * $p < 0.01$

Figure 3.

ROC curves for the different methods. There were no differences observed, and in all the area under the curve was similar.

Table 1

General characteristics of the patients.

	FC	NRFI
n	125	35
Age (years)	8.7 ± 0.4	9.2 ± 0.5
Sex (male)	58%	80% *
Duration of problem (y)	4.0 ± 0.3	4.5 ± 0.6
BM /week	4.5 ± 0.4	7.6 ± 0.7 *
Fecal incontinence (%)	62.4	100 *
Accidents/week	16.0 ± 0.4	12.8 ± 2.2
Abdominal pain (%)	33	9 *
Rectal bleeding (%)	7	0
Toilet trained (%)	70	68

*
p< 0.01.

FC=Functional constipation; NRFI= Non-retentive fecal incontinence

Table 2

Distribution of normal and abnormal scores among the different patient populations. While Barr and Leech scores were abnormal more often in children with FC. It should be noted that >20% of FC had normal scores, and > 50% of NRFI had abnormal scores. (FC= Functional constipation; NRFI= Nonretentive fecal incontinence)

	Colonic Transit	Barr	Leech	Fecal Loading
% abnormal in FC	40	79	72	63
% abnormal in NRFI	23	59	48	41
Sensitivity	0.40	0.75	0.57	0.37
Specificity	0.77	0.48	0.74	0.80

Table 3

Correlation between the different methods.

	Barr Score	Leech Score	Fecal loading	Colonic transit
Barr Score	1	.937(**)	.768(**)	.450(**)
Leech Score	.937(**)	1	.776(**)	.414(**)
Fecal loading	.768(**)	.776(**)	1	.361(**)
Colonic transit	.450(**)	.414(**)	.361(**)	1

**
p < 0.01