



UvA-DARE (Digital Academic Repository)

Constipation in (early) infancy and childhood : pathogenesis and diagnostic procedures

de Lorijn, F.

[Link to publication](#)

Citation for published version (APA):

de Lorijn, F. (2005). *Constipation in (early) infancy and childhood : pathogenesis and diagnostic procedures*. Amsterdam.

General rights

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), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: <https://uba.uva.nl/en/contact>, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.

Chapter 6

Diagnostic tests in Hirschsprung's disease; a systematic review

Fleur de Lorijn, Leontien C.M. Kremer, Johannes B. Reitsma,
and Marc A. Benninga

Submitted

Abstract

Objective

We conducted a systematic review to determine and compare the diagnostic accuracy of contrast enema, anorectal manometry, and rectal suction biopsy in infants suspected of HD.

Data sources

Articles were identified through electronic searches in Medline, Embase and Cochrane Controlled Trial Register. Searches in Pubmed were limited to articles published after 1966 and in Embase after 1980.

Study selection

Studies were included if infants underwent at least one of the following tests: contrast enema, anorectal manometry, or rectal suction biopsy, followed by full thickness biopsy and/or clinical follow-up as the reference standard.

Data extraction

Two reviewers independently assessed the methods of data collection, patient selection, blinding and prevention of verification bias, and description of the test protocol and reference standard. Data to construct 2x2 tables were abstracted for each test.

Results

Twenty-four studies met our inclusion criteria, but two studies were subsequently excluded for statistical analysis because data were missing to construct the 2x2 table. Rectal suction biopsy (14 studies for a total of 993 patients) was the most accurate test having both the highest mean sensitivity (93%, 95% CI: 88% to 95%) and mean specificity (98%, 95% CI: 95% to 99%). Sensitivity and specificity of anorectal manometry (9 studies for a total of 400 patients) was similar to that of rectal suction biopsy (91% versus 93%, $p=0.73$ and 94% versus 98%, $p=0.08$, respectively). Sensitivity and specificity of contrast enema (12 studies for a total of 425 patients) was significantly lower than that of rectal suction biopsy and anorectal manometry, with respectively 70% and 83%.

Conclusion

Rectal suction biopsy is the most accurate test in the diagnostic work-up of HD.

Introduction

Constipation is a common problem in children and has often been regarded as a trivial symptom, but it can result in substantial discomfort and disability impairing social and physical well-being. In only 10% of all children with defecation disorders, constipation is part of an organic disorder¹. About once per 5000 live births, constipation is caused by Hirschsprung's disease (HD)².

HD is a developmental disorder of the enteric nervous system (ENS) characterized by absence of ganglion cells in the myenteric and submucosal plexuses along a variable portion of the distal intestine. Literature showed that the aganglionosis is confined to rectosigmoid in 75% of patients, sigmoid, splenic flexure or transverse colon in 17% and total colon in 17% and total colon along with a short segment of the terminal ileum in 8%^{3,4}. Eighty to 90% of all cases of HD produce clinical symptoms and are diagnosed during the neonatal period. The usual presentation of HD in the neonatal period is lack of passage or delayed passage of meconium and signs and symptoms of large bowel obstruction which can lead to enterocolitis.

The diagnosis of HD is not always easy to establish. Three tests are available in the diagnostic work-up of HD. The presence of a transitional zone is the critical feature to suspect HD in contrast enema (CE) test. Anorectal manometry (ARM) assesses the recto-anal inhibition reflex (RAIR), and failure to elicit this reflex indicates HD. The third option consists of rectal suction biopsy (RSB), which shows an elevated acetyl cholinesterase activity and an aganglionosis in case of HD. However, the gold standard remains a full-thickness biopsy (FTB) of the rectum, but this is an invasive test requiring general anaesthesia.

There has been considerable debate about the most appropriate initial test for diagnosing HD, since CE, RSB and ARM can all produce false-negative and false-positive test results⁵⁻⁷. Each of these tests has both advantages and disadvantages in terms of availability, technical difficulty, radiation exposure, and invasiveness. Consequently, the choice of screening test and order of investigations for HD differs among medical centres.

We conducted a systematic review to determine and compare the diagnostic accuracy between contrast enema, anorectal manometry and rectal suction biopsy in infants suspected of HD.

Methods

Search strategy

A comprehensive literature search was performed to identify relevant publications examining the diagnostic accuracy of at least one of the following tests: CE, ARM or RSB. The Medline, Embase and the Cochrane Controlled Trial Register databases were searched using three different queries; (a) Hirschsprung [mesh or textword] AND (barium [mesh or text word] OR enema [mesh or text word]); (b) Hirschsprung [mesh or textword] AND manometry [mesh or text word]; (c) Hirschsprung [mesh or textword] AND (pathology [subheading] OR biopsy [mesh or text word] OR acetylcholinesterase [mesh or text word]). Searches in Pubmed were limited to articles published after 1966 and in Embase after 1980, and to studies involving humans. No language restriction was applied. Searches were performed in May 2003.

Selection of studies

Eligible studies had to report on the diagnostic accuracy of at least one of the following tests: contrast enema, anorectal manometry, or rectal suction biopsy in a study population suspected for HD. Staining for acetylcholinesterase (AChE) activity or staining with hematoxylin and eosin (H&E) should have been used to evaluate the rectal suction biopsy. We posed no restriction on the design of studies and included case series, case-control, cohort studies and studies in which at least one reference standard was described. For a positive index test, we defined the reference standard an aganglionosis on FTB. For a negative index test, we defined the reference standard presence of ganglion cells on FTB or on RSB, or disappearance of symptoms during FU. Studies in which we were unable to reproduce the 2 by 2 table were excluded. Articles were independently selected and reviewed by 2 authors (FdL and MB). Disagreements were resolved by consensus.

Methodological quality and data extraction

Two authors (FdL and MB) independently assessed the design of the study, blinding and prevention of verification bias, methods of data collection and patient characteristics⁸. The other two authors (L.K. and J.R.) assessed those results and disagreements were resolved by consensus. For each article we recorded the number of true-positive, false-positive, true-negative, and false-negative results for each index test that was examined.

Data synthesis

We used forrest plots to display the precision by which sensitivity and specificity had been measured in each study, and to illustrate the variation in estimates between studies. 95% confidence intervals (CI) were calculated using the exact binomial method. We used a bivariate meta-regression model to meta-analyse estimates of sensitivity and specificity^{9,10}. Rather than using a single outcome measure per study, like the diagnostic odds ratio in the Summary Receiver Operating Characteristic (SROC) approach, the bivariate model preserves the two-dimensional nature of diagnostic data by directly analysing the logit transformed sensitivity $\log(\text{sens}/(1-\text{sens}))$ and specificity $\log(\text{spec}/(1-\text{spec}))$ of each study in a single model. This model estimates and incorporates the correlation that might exist between logit sensitivity and specificity within studies due to possible differences in threshold between studies. The bivariate model uses a random effects approach for both sensitivity and specificity, allowing for heterogeneity beyond chance due to clinical or methodological differences between studies. In addition, the model acknowledges the difference in precision by which sensitivity and specificity have been measured in each study. This means that studies with a larger number of patients with the target condition receive more weight in the calculation of the summary estimate of sensitivity, while studies with more patients without the target condition are more influential in the pooling of specificity.

The model requires logit transformation of the sensitivity and specificity. A standard correction of adding 0.5 to all four cells of the 2x2 table was applied when either sensitivity or specificity was 100%. The model produces the following results: a random effect estimate of the mean sensitivity and specificity with corresponding 95% confidence intervals, the amount of between-study variation for sensitivity and specificity separately, and the strength and shape of the correlation between sensitivity and specificity. Using these results, we calculated a 95% confidence ellipse around the summary estimate of sensitivity and specificity. All the results have been transformed back (anti-logit) to the original scale, and plotted in ROC space.

Covariates indicating type of index test can be added to the model to test explicitly whether either sensitivity, or specificity, or both are different between index tests.

We performed a subgroup analysis comparing studies with a substandard verification protocol to studies with an adequate protocol. An adequate verification

protocol was defined as a study in which both, a positive index test and a negative index test, a predefined reference standard was performed. Case series in which the reference standard was performed for a positive index test were also defined as an adequate verification protocol. Unclear descriptions of reference tests were classified as a substandard verification.

The Proc Mixed procedure in SAS version 8.2 for Windows (SAS Institute Inc, Cary, NC, USA) was used to fit the various bivariate models.

Results

Literature search

We identified 587 studies from the electronic search, and after initial evaluation 54 of them were judged potentially relevant. Twenty-eight studies were excluded because no reference standard was specified and 2 studies were excluded because no AChE or H&E staining was used; the remaining 24 studies met our inclusion criteria. Twelve studies evaluated the diagnostic value of the contrast enema, 9 studies evaluated anorectal manometry and 14 studies evaluated rectal suction biopsy. The median sample size of the 24 studies was 45 patients (range 9-293). Two of these 24 articles were excluded from the statistical analysis because data were presented unclear and we were unable to produce a 2 by 2 table.

Quality assessment and data extraction

Table 1-3 lists the 24 included studies and their clinical and design characteristics. The design of studies varied: 16 studies were cohort studies, 4 studies were case series, and 1 study used a case-control design. The design was unclear in 3 studies. 14 of the 24 studies used retrospective data collection, while the method of data collection was unclear in the other 10 studies. None of the studies reported whether the readers of the index tests were blinded to the results of the other index tests. Eleven of the 12 studies evaluating CE used FTB as the reference standard to verify positive CE results. However, only 7 studies used one of our predefined reference standards to verify negative CE results. A FTB was used as reference standard for positive ARM results in 8 of the 9 studies evaluating ARM, but verification of negative ARM results was suboptimal in 3 studies. Thirteen of the 14 RSB studies used FTB as the reference standard in case of a positive RSB result. In 8 RSB studies the verification of negative test results was inadequate. Ten of the 14 RSB studies were stained with AChE, whereas 4 studies were stained with H&E. Because of the small number of studies using H&E staining results were not compared with the other index tests. In 2 studies using the AChE staining it was not possible to reproduce the 2 by 2 table and these studies were therefore not included the analysis.

Diagnostic accuracy

Figure 1 and 2 shows the sensitivity and specificity and their 95% confidence interval for each study. Confidence intervals are wide, especially for estimates of

Table 1. Evidence table showing clinical and design characteristics of individual studies. Index test is contrast enema.

First author	Year	Design	Verification		Blind
			Index +	Index -	
Davis et al. ²⁶	1972	Case Series	FTB	NA	NR
Mahboubi et al. ²⁷	1978	Case Series	FTB	NA	NR
Kekomaki et al. ²⁴	1979	Cohort	FTB in 10 infants	Unclear	NR
De Campo et al. ¹⁴	1984	Case Series	FTB	NA	NR
Lanfranchi et al. ²⁸	1984	Cohort	FTB	FU of at least 6 mo	NR
Rosenfield et al. ¹⁷	1984	Cohort	FTB	FTB	NR
Loening-Baucke et al. ¹²	1985	Cohort	Aganglionosis on RSB	Ganglioncells on RSB and FU	NR
Taxman et al. ¹³	1986	Cohort	FTB	FU or RSB with ganglioncells	NR
Smith et al. ²⁹	1991	Case Series	FTB	NA	NR
O'Donovan et al. ³⁰	1996	Cohort	FTB	Combination of ARM + CE + FU	NR
Osatakul et al. ⁶	1999	Cohort	FTB	FU	NR
Reid et al. ³¹	2000	Cohort	FTB	ARM and/or ganglioncells on RSB	NR

Table 2. Evidence table showing clinical and design characteristics of individual studies. Index test is anorectal manometry.

First author	Year	Design	Verification		Blind
			Index +	Index -	
Frenckner et al. ³²	1978	Cohort	FTB	FU of at least 4 mo.	NR
Mahboubi et al. ²⁷	1978	Case Series	FTB	FTB	NR
Iwai et al. ²²	1979	Case Control	FTB	FU and negative CE and 30% FTB	NR
Kekomaki et al. ²⁴	1979	Cohort	FTB in 10 infants	NR	NR
Lanfranchi et al. ²⁸	1984	Cohort	FTB	FU of at least 6 mo.	NR
Loening-Baucke et al. ¹²	1985	Cohort on RSB	Aganglionosis on RSB and FU	Ganglioncells on	NR
Yaxoing et al. ³³	1986	Unclear	FTB	NR	NR
Emir et al. ⁵	1997	Cohort	FTB	38% FTB or FU at least 4 mo.	NR
Osatakul et al. ⁶	1999	Cohort	FTB	FU	NR

NR=not reported; FTB=full thickness biopsy; TP=true positive test results; FP=false positive test results; FN=false negative test results; TN=true negative test results; NA=not applicable; CE=contrast enema, ARM=anorectal manometry, RSB=rectal suction biopsy

Data collection	Population		TP	FP	TN	FN	Inconclusive
	Age	% Boys					
Retro	NR	79	22	NA	NA	6	0
Retro	25/27≤2yrs	67	22	NA	NA	3	0
Retro	Range 2wks-11yrs	Unclear	10	5	45	0	0
Retro	Unclear	54	2	NA	NA	11	0
Retro	Mean 4.8yrs (17dys-16yrs)	62	13	0	15	6	0
Retro	Unclear	Unclear	29	4	16	13	0
Unclear	Mean 14 days (2-28 days)	52	3	7	14	1	0
Retro	Range 1day-1yr	NR	12	9	29	3	0
Retro	Median 2-5 days	NR	13	NA	NA	4	0
Retro	Range 2dys-9mo	61	13	NR	NR	5	0
Unclear	Range 0.2-120 mo	64	21	0	5	10	0
Retro	Mean 3 yrs (4wks-15yrs)	56	5	1	47	1	0

Data collection	Population		TP	FP	TN	FN	Inconclusive
	Age	% Boys					
Unclear	Median 5 mo (1mo-12mo)	NR	3	0	5	0	1
Retro	25/27≤2yrs	67	22	NA	NA	3	0
Unclear	Range 4mo.-19yrs	NR	9	0	10	0	0
Retro	Range 2wks-11yrs	Unclear	8	1	47	1	3
Retro	Mean 4.8yrs (17dys-16yrs)	62	17	0	15	2	0
Unclear	Mean 14 days (2-28 days)	52	3	1	20	1	0
Unclear	Range 30dys-12yrs	59	63	5	58	0	0
Unclear	Mean 30,3 dys (2dys-3mo)	63	22	1	32	1	0
Unclear	Range 0.2-120 mo	64	31	2	12	0	0

Table 3. Evidence table showing clinical and design characteristics of individual studies. Index test is rectal suction biopsy.

First author,	year	Design	Verification		Blind
			Index +	Index -	
Campbell et al. ³⁴	1969	Unclear	FTB	NR	NR
Kekomaki et al. ²⁴	1979	Cohort	FTB in 10 infants	NR	NR
Hamoudi et al. ³⁶	1982	Unclear	FTB	NR	NR
Huntley et al. ²³	1982	Cohort	FTB	Ganglioncells on RSB	NR
Barr et al. ³⁷	1985	Cohort	FTB	FU of 18 months	NR
Loening-Baucke et al. ¹²	1985	Cohort	Aganglionosis on RSB	Ganglioncells on RSB and FU of at least 6 months	NR
Kurer et al. ³⁸	1986	Cohort	FTB	NR	NR
Polley et al. ³⁹	1986	Cohort	FTB	RSB with ganglioncells	NR
Taxman et al. ¹³	1986	Cohort	FTB	FU or RSB with ganglioncells	NR
Yaxoing et al. ³³	1986	Unclear	FTB	NR	NR
Bonham et al. ⁴⁰	1987	Cohort	FTB	NR	NR
Chen et al. ²⁵	1987	Cohort	FTB	FTB or FU of at least 1 mo.	NR
Smith et al. ²⁹	1991	Case Series	FTB	NR	NR
Park et al. ⁴¹	1992	Cohort	FTB in 13 infants	NR	NR

NR=not reported; FTB=full thickness biopsy; TP=true positive test results; FP=false positive test results; FN=false negative test results; TN=true negative test results; NA=not applicable; CE=contrast enema, ARM=anorectal manometry, RSB=rectal suction biopsy

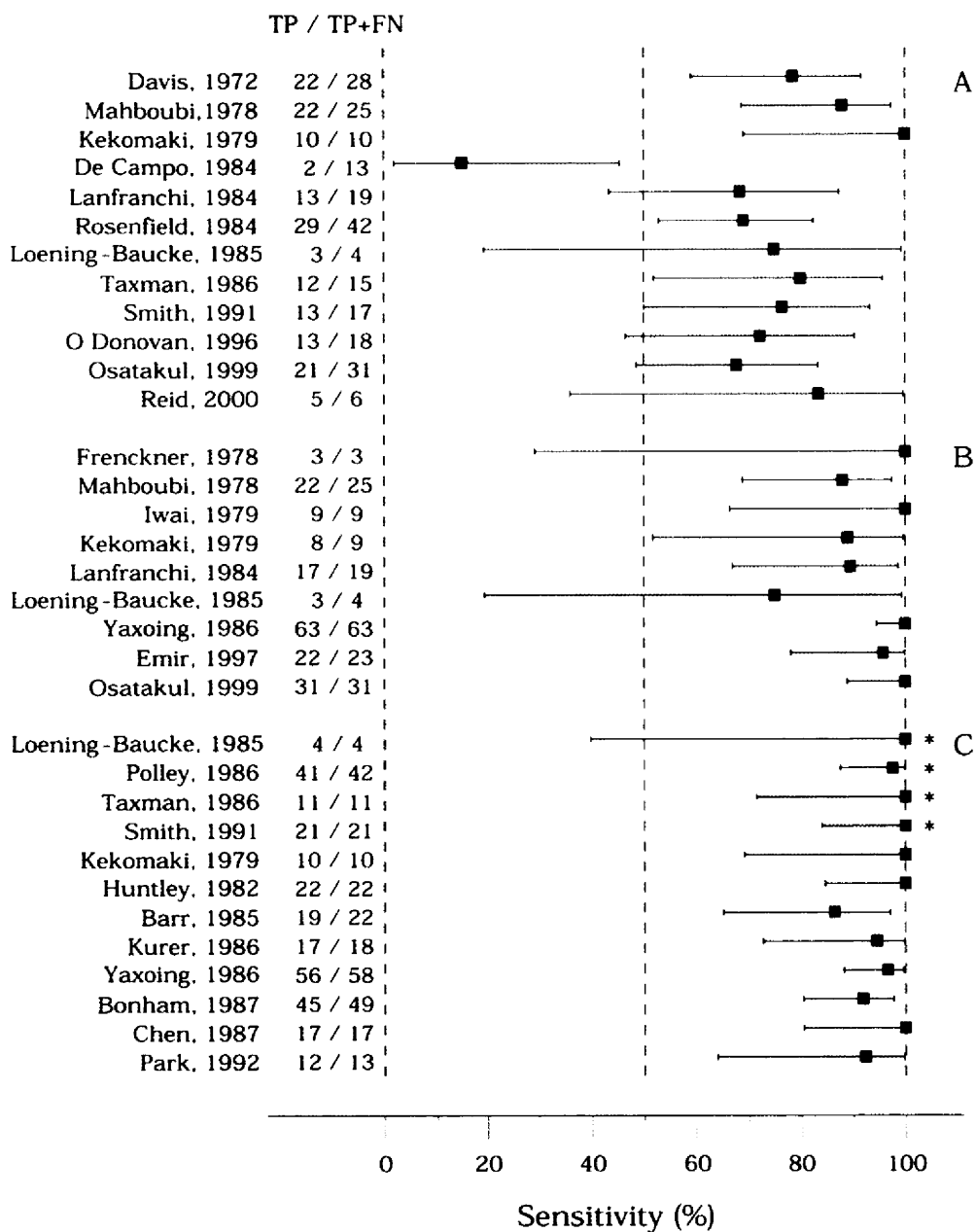
Table 4. Summary estimates for sensitivity and specificity from the bivariate model. Comparison between contrast enema, anorectal manometry and rectal suction biopsy stained for AChE activity.

Index test	Parameter	
	Mean sensitivity(95% CI)	Mean specificity(95% CI)
Contrast enema (CE)	0.70 (0.64 to 0.76)	0.83 (0.74 to 0.90)
Anorectal manometry (ARM)	0.91 (0.85 to 0.95)	0.94 (0.89 to 0.97)
Rectal suction biopsy (RSB)	0.93 (0.88 to 0.95)	0.98 (0.95 to 0.99)
P-value CE vs. ARM	<0.0001	0.014
P-value CE vs. RSB	<0.0001	<0.0001
P-value ARM vs. RSB	0.730	0.085

CI = confidence interval.

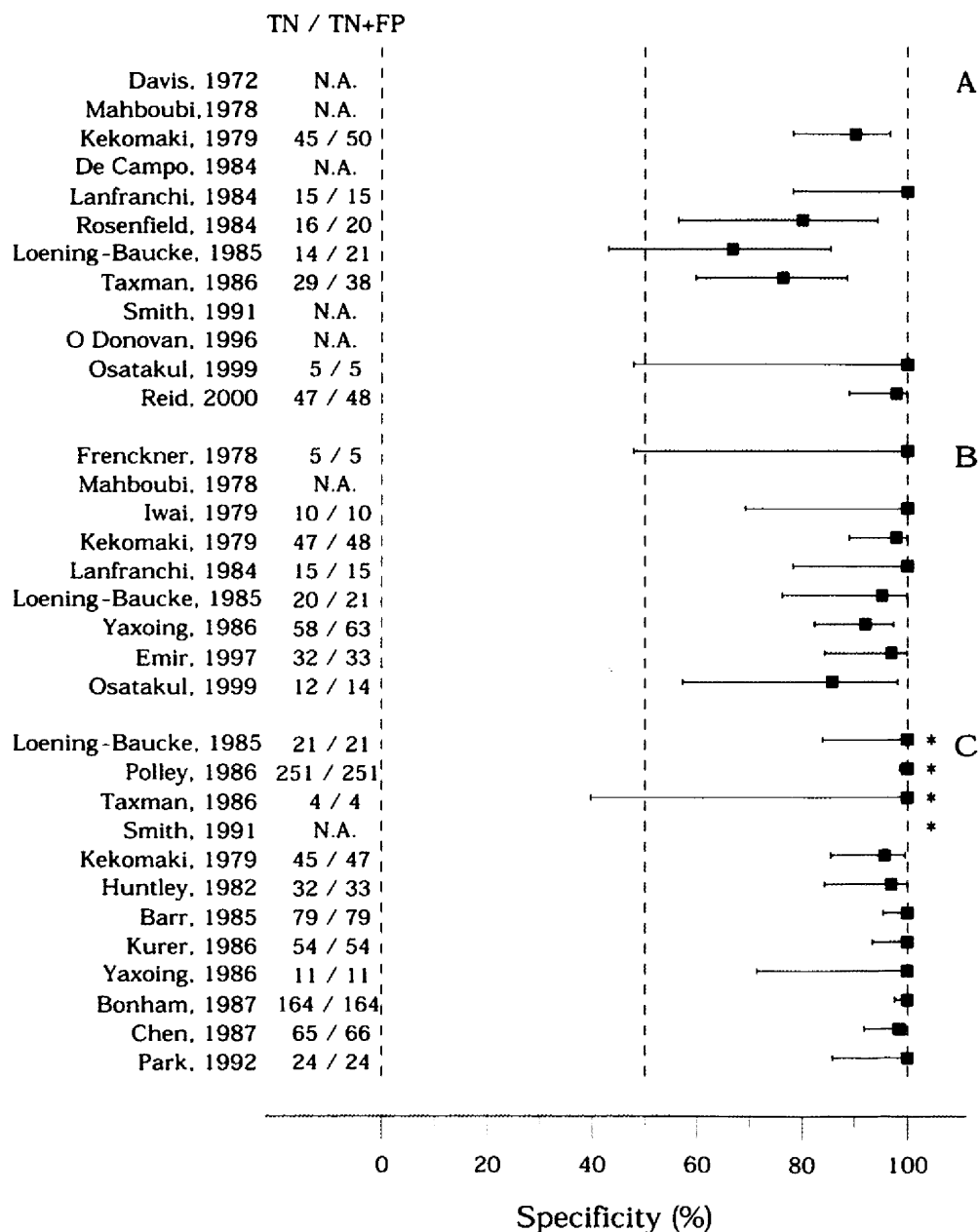
Data collection	Population		TP	FP	TN	FN	Inconclusive
	Age	% Boys					
Unclear	Unclear	NR	NR	NR	NR	NR	NR
Retro	1day-12yrs						
	Range	Unclear	10	2	45	0	3
	2wks-11yrs						
Unclear	NR	NR	NR	NR	NR	NR	NR
Retro	Range	NR	22	1	32	0	3
	4dys-12yrs						
Retro	NR	NR	19	0	79	3	0
Unclear	Mean 14 days	52	4	0	21	0	0
	(2-28 days)						
Unclear	Range few	NR	17	0	54	1	0
	weeks-17yrs						
Retro	Mean 14.4 mo	NR	41	0	251	1	0
Retro	Range	NR	11	0	40	0	7
	1day-1yr						
Unclear	Range	59	56	0	11	2	0
	30dys-12yrs						
Retro	Range	NR	45	0	164	4	0
	2dys-14yrs						
Unclear	Range	NR	17	1	65	0	0
	3dys-10yrs						
	(2dys-15yrs)						
Retro	Median	NR	21	NA	NA	0	3
	2-5 dys						
Unclear	Range	Unclear	12	0	24	1	0
	3 dys-17yrs						

Figure 1. Forrest plot of sensitivity in studies examining the accuracy of contrast enema (A), anorectal manometry (B), and rectal suction biopsy (C). TP=true positive test result; FN=false negative test result.



*RSB studies stained with H&E.

Figure 2. Forrest plot of specificity in studies examining the accuracy of contrast enema (A), anorectal manometry (B), and rectal suction biopsy (C). TN=true negative test result; FP=false positive test result. NA= not applicable, case series that only included patients with Hirschsprung.



*RSB studies stained with H&E.

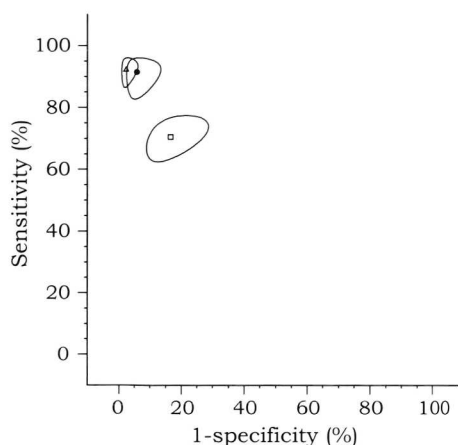


Figure 3. Bivariate random summary estimates of sensitivity and specificity for each of the three index tests (square=contrast enema, dot=anorectal manometry, triangle=rectal suction biopsy stained for AChE activity) and the corresponding 95% confidence ellipse around these means.

sensitivity because the number of patients with HD is low in most studies. The estimates of mean sensitivity and specificity of the three index tests are given in table 4. The joint confidence interval around the mean sensitivity and specificity is depicted in figure 3. This figure clearly shows that the accuracy of CE is substantially poorer compared to the other tests. Formal statistical testing for differences between sensitivity and specificity reveal that RSB stained for AChE activity was the most accurate test having both the highest mean sensitivity (93%, 95% CI: 88% to 95%) and mean specificity (98%, 95% CI: 95% to 99%). The RSB studies stained for H&E showed a comparable sensitivity and specificity (96% and 98% respectively). Sensitivity and specificity of anorectal manometry (9 studies for a total of 400 patients) was similar to that of rectal suction biopsy stained for AChE activity (91% vs 93%, $p=0.73$ and 94%, vs 98%; $p=0.08$, respectively). Sensitivity and specificity of CE was significantly lower than that of RSB stained for AChE activity and ARM, with respectively 70% and 83%.

Subgroup analyses showed that there was no significant difference in either sensitivity or specificity between studies with a low quality and high quality verification protocol. However, a trend towards an exaggerated estimate of sensitivity was found in low ($n=2$) versus high quality ($n=10$) studies evaluating CE (mean sensitivity of low quality 86% vs. high quality 70%), but this difference was not statistical significant ($p=0.33$).

Inconclusive tests were most frequently seen in studies evaluating RSB; 4 studies showed 16 inconclusive tests and in 2 studies the number of inconclusive tests was unclear. ARM showed in 2 studies 4 inconclusive tests and CE did not show inconclusive tests.

Discussion

Our results show that rectal suction biopsy (RSB) stained for AChE activity is the most accurate test in the diagnostic work-up of HD. Mean sensitivity of anorectal manometry (ARM) was similar to that of RSB, but its specificity was significantly lower. Mean sensitivity and specificity of contrast enema (CE) was significantly lower than that of RSB and ARM.

To our knowledge, this is the first systematic review evaluating the accuracy of tests for the diagnosis of HD. Since it is not clear what test is the most accurate one in the diagnostic work-up of HD, different approaches are used in different hospitals. Only one report in this systematic review compared all three possible tests in a single study¹¹. Therefore, we used indirect comparisons to examine differences in sensitivities and specificities between the three tests.

All studies included in this systematic review had their limitations. The way patient were selected was poorly described in all studies. Therefore, it was unclear whether all consecutive series of patients suspected of HD was included or not. Furthermore, a description of the clinical symptoms of patients suspected of HD was lacking. For most studies it was unclear whether the readers of the index tests were blinded to the results of the other index tests or reference standard. Another drawback is that all studies provided sparse details about the thoroughness of their clinical follow-up. Essential information like the duration of follow-up, the number of patients lost to follow-up, and the way in which patients were contacted, was missing. All studies retrospectively collected their data. The disadvantage of collecting data retrospectively is: uncertainty about inclusion criteria, that data-collection is not uniform, evaluating known test results and that quality of follow-up is limited. In almost all studies (23 studies) included in this systematic review a FTB was performed in those patients with a positive index test. However, several studies did not verify patients with negative test results and classified these patients as true negatives. Patients with a false negative test result may stay undetected if follow-up is not adequate or if patients with persistent symptoms return to a different hospital. In that case, both sensitivity and specificity will be overestimated, although the relative effect will be greater on sensitivity because of the lower number of patients with HD.

In this systematic review, CE had the lowest sensitivity compared to RSB and ARM (70% vs 93% and 91% respectively), indicating that false-negative test results

were more common in CE. Several explanations have been described for false-negative results of CE. Especially in young infants suspected of HD, a transition zone is difficult to demonstrate¹². Furthermore, De Campo et al. demonstrated a normal-calibre colon in 75% of children with total aganglionosis¹³. Rectal wash-outs and even digital rectal examinations may also lead to false-negative test results¹⁴. RSB and ARM showed in rare cases false-negative test results. It has been suggested that false-negative test results of ARM occur because of displacement of the transducer probe or as a consequence of relaxation of the external anal sphincter rather than the internal anal sphincter¹⁵. Possible causes for false-negative test results in RSB are: variability in the biopsy site (correct site is 2-3 cm above the mucocutaneous junction), biopsy material that is too superficial and does not contain muscularis mucosa, biopsy material with bleeding artefacts, the technical variations in performance of the stain and the experience of individual pathologists. Furthermore, in neonates a false-negative reaction may occur possibly due to immaturity of the enzyme system¹⁴.

A high specificity is also desirable for tests in the work-up of patients suspected for HD, otherwise too many patients unnecessarily have to undergo the invasive definitive test because of false-positive test results. False-positive test results were also most frequent in CE. Rosenfield et al. described 4 newborn infants with a meconium plug syndrome leading to a transition zone at the splenic flexure, that mimics the finding in HD¹⁶. Other CE studies did not provide a good explanation for the occurrence of false-positive test results. The literature contains conflicting views whether the accuracy of anorectal manometry is lower in neonates. Some studies suggest that failure of the reflex might be due to physiologic immaturity of anorectal function^{17,18}. Nevertheless, we demonstrated that term and premature infants older than 26 weeks' postmenstrual age have a well developed RAIR upon rectal distension^{19,20}. The false-positive test results in ARM might be due to technical problems such as an air leak in the circuit or insufficient inflation of the balloon²¹. False-positive results in RSB are rare. Only three studies showed patients with false-positive test results²²⁻²⁴. The reason for this is unclear.

The highest frequency of inconclusive tests results were seen in studies evaluating RSB, 4 of the 14 studies showed 16 inconclusive tests. The explanation is that biopsies are taken too superficially (containing insufficient muscularis mucosa) for proper evaluation. In contrast, CE did not show inconclusive tests.

In this systematic review, CE showed a wide range in sensitivity and specificity. The wide range in sensitivity and specificity values might be due to difference in quality of studies and/or variation due to chance because of low number of included patients. CE showed 1 outlier with a sensitivity of only 15%¹³. The reason of this low sensitivity can be the inclusion of only 13 patients, all with total aganglionosis which often leads to false-negative test results on CE. Specificity of CE studies showed also a wide range and could be calculated in only 7 studies. The range of sensitivity and specificity in RSB studies is small. This is also true for the sensitivity and specificity in ARM studies.

In conclusion, this systematic review shows that rectal suction biopsy stained for AChE activity is the most accurate test in the diagnostic work-up of patients suspected of HD. Since verification of index tests was not optimal in all studies included in this systematic review and all studies showed methodological limitations, the value of the accuracy of these tests need to be confirmed. Large blinded prospective studies with well defined inclusion criteria and long-term clinical follow up should be performed to evaluate the accuracy of the different diagnostic tests simultaneously, in infants suspected of Hirschsprung's disease.

References

1. Loening-Baucke V. Chronic constipation in children. *Gastroenterology* 1993;105:1557-64.
2. Puri P. Hirschsprung's disease: Clinical Generalities. In Holschneider AM, Puri P editors. *Hirschsprung's disease and allied disorders*. 2000;129-35.
3. Ikeda K, Goto S. Diagnosis and treatment of Hirschsprung's disease in Japan. An analysis of 1628 patients. *Ann Surg* 1984;199:400-5.
4. Kleinhaus S, Boley SJ, Sheran M et al. Hirschsprung's disease - a survey of the members of the Surgical Section of the American Academy of Pediatrics. *J Pediatr Surg* 1979;14:588-97.
5. Emir H, Akman M, Sarimurat N et al. Anorectal manometry during the neonatal period: its specificity in the diagnosis of Hirschsprung's disease. *Eur J Pediatr Surg* 1999;9:101-3.
6. Osatakul S, Patrapinyokul S, Osatakul N. The diagnostic value of anorectal manometry as a screening test for Hirschsprung's disease. *J Med Assoc Thai* 1999;82:1100-5.
7. De Campo JF, Mayne V, Boldt DW et al. Radiological findings in total aganglionosis coli. *Pediatr Radiol* 1984;14:205-9.
8. Bossuyt P, Reitsma J, Bruns D et al. The STARD statement for reporting studies of diagnostic accuracy: Explanation and elaboration. *Ann Intern Med* 2003;138:w1-12.
9. van Houwelingen HC, Zwinderman KH, Stijnen T. A bivariate approach to meta-analysis. *Stat Med* 1993;12:2273-84.
10. van Houwelingen HC, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. *Stat Med* 2002;21:589-624.
11. Loening-Baucke V, Pringle KC, Ekwo EE. Anorectal manometry for the exclusion of Hirschsprung's disease in neonates. *J Pediatr Gastroenterol Nutr* 1985;4:596-603.
12. Taxman TL, Yulish BS, Rothstein FC. How useful is the barium enema in the diagnosis of infantile Hirschsprung's disease? *Am J Dis Child* 1986;140:881-4.
13. De Campo JF, Mayne V, Boldt DW et al. Radiological findings in total aganglionosis coli. *Pediatr Radiol* 1984;14:205-9.
14. Blake NS. Diagnosis of Hirschsprung's Disease and Allied Disorders. In Holschneider AM, Puri P editors. *Hirschsprung's Disease and allied disorders*. 2000;223-90.
15. Aaronson I, Nixon HH. A clinical evaluation of anorectal pressure studies in the diagnosis of Hirschsprung's disease. *Gut* 1972;13:138-46.
16. Rosenfield NS, Ablow RC, Markowitz RI et al. Hirschsprung disease: accuracy of the barium enema examination. *Radiology* 1984;150:393-400.

17. Holschneider AM, Kellner E, Streibl P et al. The development of anorectal continence and its significance in the diagnosis of Hirschsprung's disease. *J Pediatr Surg* 1976;11:151-6.
18. Ito Y, Donahoe PK, Hendren WH. Maturation of the rectoanal response in premature and perinatal infants. *J Pediatr Surg* 1977;12:477-82.
19. Benninga MA, Omari TI, Haslam RR et al. Characterization of anorectal pressure and the anorectal inhibitory reflex in healthy preterm and term infants. *J Pediatr* 2001;139:233-7.
20. de Lorijn F, Omari TI, Kok JH et al. Maturation of the rectoanal inhibitory reflex in very premature infants. *J Pediatr* 2003;143:630-3.
21. Iwai N, Ogita S, Kida M et al. A manometric assessment of anorectal pressures and its significance in the diagnosis of Hirschsprung's disease and idiopathic megacolon. *Jpn J Surg* 1979;9:234-40.
22. Huntley CC, Shaffner LD, Challa VR et al. Histochemical diagnosis of Hirschsprung disease. *Pediatrics* 1982;69:755-61.
23. Kekomaki M, Rapola J, Louhimo I. Diagnosis of Hirschsprung's disease. *Acta Paediatr Scand* 1979;68:893-7.
24. Chen CL, Hsu HC, Chen CC et al. Acetylcholinesterase activity in rectal suction biopsy for the diagnosis of Hirschsprung's disease. *Taiwan Yi Xue Hui Za Zhi* 1987;86:723-7.
25. Davis PW, Foster DB. Hirschsprung's disease. A clinical review. *Br J Surg* 1972;59:19-26.
26. Mahboubi S, Schnauffer L. The barium-enema examination and rectal manometry in Hirschsprung disease. *Radiology* 1978;130:643-7.
27. Lanfranchi GA, Bazzocchi G, Federici S et al. Anorectal manometry in the diagnosis of Hirschsprung's disease - comparison with clinical and radiological criteria. *Am J Gastroenterol* 1984;79:270-5.
28. Smith GHH, Cass D. Infantile Hirschsprung's disease - Is a barium enema useful? *Pediatr Surg Int* 1991;6:318-21.
29. O'Donovan AN, Habra G, Somers S et al. Diagnosis of Hirschsprung's disease. *AJR Am J Roentgenol* 1996;167:517-20.
30. Reid JR, Buonomo C, Moreira C et al. The barium enema in constipation: Comparison with rectal manometry and biopsy to exclude Hirschsprung's disease after the neonatal period. *Pediatr Radiology* 2000;30:681-4.
31. Freneckner B. Ano-rectal manometry in the diagnosis of Hirschsprung's disease in infants. *Acta Paediatr Scand* 1978;67:187-92.
32. Yaxiong S, Chengren S. Clinical evaluation of diagnostic methods for Hirschsprung's disease. *Pediatr Surg Int* 1986;1:218-22.

33. Campbell PE, Noblett HR. Experience with rectal suction biopsy in the diagnosis of Hirschsprung's disease. *J Pediatr Surg* 1969;4:410-5.
34. Hamoudi AB, Reiner CB, Boles ET et al. Acetylthiocholinesterase staining activity of rectal mucosa. Its use in the diagnosis of Hirschsprung's disease. *Arch Pathol Lab Med* 1982;106:670-2.
35. Barr LC, Booth J, Filipe MI et al. Clinical evaluation of the histochemical diagnosis of Hirschsprung's disease. *Gut* 1985;26:393-9.
36. Kurer MH, Lawson JO, Pambakian H. Suction biopsy in Hirschsprung's disease. *Arch Dis Child* 1986;61:83-4.
37. Polley Jr TZ, Coran AG, Heidelberger KP et al. Suction rectal biopsy in the diagnosis of Hirschsprung's disease and chronic constipation. *Pediatr Surg Int* 1986;1:84-9.
38. Bonham JR, Dale G, Scott DJ et al. A 7-year study of the diagnostic value of rectal mucosal acetylcholinesterase measurement in Hirschsprung's disease. *J Pediatr Surg* 1987;22:150-2.
39. Park WH, Choi SO, Kwon KY et al. Acetylcholinesterase histochemistry of rectal suction biopsies in the diagnosis of Hirschsprung's disease. *J Korean Med Sci* 1992;7:353-9.