Cochrane Library

Cochrane Database of Systematic Reviews

Imaging modalities for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome (Review)

van Gruting IMA, Stankiewicz A, Thakar R, Santoro GA, IntHout J, Sultan AH

van Gruting IM A, Stankiewicz A, Thakar R, Santoro GA, IntHout J, Sultan AH.
Imaging modalities for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome.
Cochrane Database of Systematic Reviews 2021, Issue 9. Art. No.: CD011482.
DOI: 10.1002/14651858.CD011482.pub2.

www.cochranelibrary.com

WILEY

TABLE OF CONTENTS

[Diagnostic Test Accuracy Review]

Imaging modalities for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome

Isabelle M A van Gruting[1], Aleksandra Stankiewicz[2], Ranee Thakar[3], Giulio A Santoro[4], Joanna IntHout[5], Abdul H Sultan[3]

[1]Department of Obstetrics and Gynaecology, Croydon University Hospital NHS Trust, Croydon, United Kingdom. [2]Department of Radiology, Croydon University Hospital, Croydon, UK. [3]Department of Obstetrics and Gynaecology, Croydon University Hospital NHS Trust, Croydon, UK. [4]Section of Anal Physiology and Ultrasound, Department of Surgery, Regional Hospital, Treviso, Italy. [5]Radboud Institute for Health Sciences, Radboud university medical center, Nijmegen, Netherlands

Contact: Isabelle M A van Gruting, isabellevangruting@gmail.com.

Editorial group: Cochrane Colorectal Group.
Publication status and date: Edited (no change to conclusions), published in Issue 9, 2021.

Citation: van Gruting IM A, Stankiewicz A, Thakar R, Santoro GA, IntHout J, Sultan AH. Imaging modalities for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome. *Cochrane Database of Systematic Reviews* 2021, Issue 9. Art. No.: CD011482. DOI: 10.1002/14651858.CD011482.pub2.

ABSTRACT

Background

Obstructed defaecation syndrome (ODS) is difficulty in evacuating stools, requiring straining efforts at defaecation, having the sensation of incomplete evacuation, or the need to manually assist defaecation. This is due to a physical blockage of the faecal stream during defaecation attempts, caused by rectocele, enterocele, intussusception, anismus or pelvic floor descent. Evacuation proctography (EP) is the most common imaging technique for diagnosis of posterior pelvic floor disorders. It has been regarded as the reference standard because of extensive experience, although it has been proven not to have perfect accuracy. Moreover, EP is invasive, embarrassing and uses ionising radiation. Alternative imaging techniques addressing these issues have been developed and assessed for their accuracy. Because of varying results, leading to a lack of consensus, a systematic review and meta-analysis of the literature are required.

Objectives

To determine the diagnostic test accuracy of EP, dynamic magnetic resonance imaging (MRI) and pelvic floor ultrasound for the detection of posterior pelvic floor disorders in women with ODS, using latent class analysis in the absence of a reference standard, and to assess whether MRI or ultrasound could replace EP. The secondary objective was to investigate differences in diagnostic test accuracy in relation to the use of rectal contrast, evacuation phase, patient position and cut-off values, which could influence test outcome.

Search methods

We ran an electronic search on 18 December 2019 in the Cochrane Library, MEDLINE, Embase, SCI, CINAHL and CPCI. Reference list, Google scholar. We also searched WHO ICTRP and clinicaltrials.gov for eligible articles. Two review authors conducted title and abstract screening and full-text assessment, resolving disagreements with a third review author.

Selection criteria

Diagnostic test accuracy and cohort studies were eligible for inclusion if they evaluated the test accuracy of EP, and MRI or pelvic floor ultrasound, or both, for the detection of posterior pelvic floor disorders in women with ODS. We excluded case-control studies. If studies partially met the inclusion criteria, we contacted the authors for additional information.

Data collection and analysis

Two review authors performed data extraction, including study characteristics, 'Risk-of-bias' assessment, sources of heterogeneity and test accuracy results. We excluded studies if test accuracy data could not be retrieved despite all efforts. We performed meta-analysis

using Bayesian hierarchical latent class analysis. For the index test to qualify as a replacement test for EP, both sensitivity and specificity should be similar or higher than the historic reference standard (EP), and for a triage test either specificity or sensitivity should be similar or higher. We conducted heterogeneity analysis assessing the effect of different test conditions on test accuracy. We ran sensitivity analyses by excluding studies with high risk of bias, with concerns about applicability, or those published before 2010. We assessed the overall quality of evidence (QoE) according to GRADE.

Main results

Thirty-nine studies covering 2483 participants were included into the meta-analyses. We produced pooled estimates of sensitivity and specificity for all index tests for each target condition. Findings of the sensitivity analyses were consistent with the main analysis.

Sensitivity of EP for diagnosis of rectocele was 98% (credible interval (CrI)94%-99%), enterocele 91%(CrI 83%-97%), intussusception 89%(CrI 79%-96%) and pelvic floor descent 98%(CrI 93%-100%); specificity for enterocele was 96%(CrI 93%-99%), intussusception 92%(CrI 86%-97%) and anismus 97%(CrI 94%-99%), all with high QoE. Moderate to low QoE showed a sensitivity for anismus of 80%(CrI 63%-94%), and specificity for rectocele of 78%(CrI 63%-90%) and pelvic floor descent 83%(CrI 59%-96%).

Specificity of MRI for diagnosis of rectocele was 90% (CrI 79%-97%), enterocele 99% (CrI 96%-100%) and intussusception 97% (CrI 88%-100%), meeting the criteria for a triage test with high QoE. MRI did not meet the criteria to replace EP. Heterogeneity analysis showed that sensitivity of MRI performed with evacuation phase was higher than without for rectocele (94%, CrI 87%-98%) versus 65%, CrI 52% to 89%, and enterocele (87%, CrI 74%-95% versus 62%, CrI 51%-88%), and sensitivity of MRI without evacuation phase was significantly lower than EP.

Specificity of transperineal ultrasound (TPUS) for diagnosis of rectocele was 89% (CrI 81%-96%), enterocele 98% (CrI 95%-100%) and intussusception 96% (CrI 91%-99%); sensitivity for anismus was 92% (CrI 72%-98%), meeting the criteria for a triage test with high QoE. TPUS did not meet the criteria to replace EP. Heterogeneity analysis showed that sensitivity of TPUS performed with rectal contrast was not significantly higher than without for rectocele(92%, CrI 69%-99% versus 81%, CrI 58%-95%), enterocele (90%, CrI 71%-99% versus 67%, CrI 51%-90%) and intussusception (90%, CrI 69%-98% versus 61%, CrI 51%-86%), and was lower than EP.

Specificity of endovaginal ultrasound (EVUS) for diagnosis of rectocele was 76% (CrI 54%-93%), enterocele 97% (CrI 80%-99%) and intussusception 93% (CrI 72%-99%); sensitivity for anismus was 84% (CrI 59%-96%), meeting the criteria for a triage test with very low to moderate QoE. EVUS did not meet the criteria to replace EP.

Specificity of dynamic anal endosonography (DAE) for diagnosis of rectocele was 88% (CrI 62%-99%), enterocele 97% (CrI 75%-100%) and intussusception 93% (CrI 65%-99%), meeting the criteria for a triage test with very low to moderate QoE. DAE did not meet the criteria to replace EP.

Echodefaecography (EDF) had a sensitivity of 89% (CrI 65%-98%) and specificity of 92% (CrI 72%-99%) for intussusception, meeting the criteria to replace EP but with very low QoE. Specificity of EDF for diagnosis of rectocele was 89% (CrI 60%-99%) and for enterocele 97% (CrI 87%-100%); sensitivity for anismus was 87% (CrI 72%-96%), meeting the criteria for a triage test with low to very low QoE.

Authors' conclusions

In a population of women with symptoms of ODS, none of the imaging techniques met the criteria to replace EP. MRI and TPUS met the criteria of a triage test, as a positive test confirms diagnosis of rectocele, enterocele and intussusception, and a negative test rules out diagnosis of anismus. An evacuation phase increased sensitivity of MRI. Rectal contrast did not increase sensitivity of TPUS. QoE of EVUS, DAE and EDF was too low to draw conclusions. More well-designed studies are required to define their role in the diagnostic pathway of ODS.

PLAIN LANGUAGE SUMMARY

Is evacuation proctogram still the reference standard for diagnosis of posterior pelvic floor disorders in women with obstructed defaecation syndrome?

The issue

Obstructed defaecation syndrome is a sensation of obstruction during attempts to empty the bowel, a feeling of incomplete bowel emptying, or the need to use a finger to splint the perineum/vagina or insert into the rectum to remove stool. This can cause embarrassment and frustration, leading to an adverse effect on quality of life. Different imaging techniques exist to examine women with these symptoms. The most commonly performed technique currently used is called evacuation proctography (EP). This test can cause embarrassment, as it requires the woman to have a large amount of a porridge-like substance introduced via the back passage and then she has to sit on a commode and open her bowels whilst X-ray images are being taken by the radiologist.

Why is this review important?

Other imaging techniques to assess women with these symptoms are available, and most of them are less embarrassing. However, it remains unclear how good these imaging techniques are in diagnosing the conditions that cause these symptoms. To be able to provide evidence for potential use of these less embarrassing imaging techniques, existing data of previously-published studies reporting the accuracy (the ability to detect and exclude a specific disorder) of these imaging techniques need to be analysed.

How was this review conducted?

We searched the available literature on 18 December 2019. We selected studies that assessed the performance of magnetic resonance imaging (MRI) or pelvic floor ultrasound, or both, and EP in women with symptoms of obstructed defaecation. We assessed the quality of the included studies, as well as possible sources that might influence the performance of imaging techniques. We conducted statistical analysis by assessing all available imaging techniques equally, in the absence of a reference standard, to calculate the test accuracy of all imaging techniques under evaluation.

What are the findings?

We included 39 studies covering 2483 women in the meta-analysis. EP was found to have the highest ability to correctly detect most conditions causing symptoms of obstructed defaecation; none of the other diagnostic tests met the criteria to replace EP. MRI and transperineal ultrasound (TPUS) met the criteria for a triage test. They are better able to correctly identify healthy patients than EP. This means that a positive test result suggests the presence of the disease, as the test rarely gives positive results in healthy women, and avoids further testing. The results of the other ultrasound techniques were of too low a quality of evidence to draw conclusions.

What does this mean?

In a population of women seeking help for their symptoms of obstructed defaecation, EP remains the test of choice. MRI and TPUS can be used for the initial assessment of women with obstructed defaecation as a screening test. TPUS or MRI could therefore potentially reduce the number of women having to undergo EP.

SUMMARY OF FINDINGS

Summary of findings 1. Summary of findings: Diagnostic test accuracy of imaging for diagnosis of posterior pelvic floor disorders in women with ODS

Review question	**What is the diagnostic test accuracy of imaging techniques for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome?**
Importance	To assess diagnostic test accuracy of imaging techniques to find an accurate, but less invasive and more patient-friendly test that could potentially replace the use of EP for the assessment of women with symptoms of obstructed defaecation syndrome (ODS)
Population[1]	Women with symptoms of ODS
Setting	Secondary and tertiary gynaecology or colorectal surgery outpatient clinics
Prior testing	History
Index tests[2]	Evacuation proctogram, magnetic resonance imaging, transperineal ultrasound, endovaginal ultrasound, dynamic anorectal ultrasound, echodefaecography
Reference standard[3]	No reference standard is available; evacuation proctography was the first available test but does not have perfect test accuracy. Statistical analysis with latent class analysis[4] was used as alternative in the absence of a reference standard
Target conditions	Rectocele, enterocele, intussusception, anismus, pelvic floor descent
Criteria for test purpose	Replacement test: both sensitivity and specificity are similar or higher than the historic reference standard EP (probability > 0.40 for sensitivity and specificity). SpIN triage test (high Specificity rules-IN the diagnosis): specificity is similar or higher than EP (probability > 0.40 for specificity, no restrictions for sensitivity). SnOUT triage test (high Sensitivity rules-OUT the diagnosis): sensitivity is similar or higher than EP (probability > 0.40 for sensitivity, no restrictions for specificity).

Test	Numer of participants	Pooled prevalence % (95% CrI)	Pooled estimate sensitivity in % (95% CrI)	Quality of the evidence (GRADE)	Pooled estimate specificity in % (95% CrI)	Quality of the evidence (GRADE)	Natural frequencies expressed in a cohort of 1000 *Based on pooled estimated prevalence by target condition*	Meets criteria for triage/replacement test and implication [7]

	(studies) [5]			for sensitivity [6]		for specificity [6]	True positives *correctly present*	False positives *over-diagnosis*	False negatives *missed*	True negatives *correctly absent*	
Rectocele											
EP	1737 (34)	58.9 (51.3 to 67.8)	97.5 (93.7 to 99.3)	⊕⊕⊕⊕ High[a]	77.8 (63.5 to 90.2)	⊕⊕⊕⊖ Moderate[a]	574	91	15	320	N/A
MRI	659 (19)		94.3 (85.9 to 98.4)	⊕⊕⊕⊕ High[b]	90.3 (78.5 to 97.4)	⊕⊕⊕⊕ High[b]	555	40	34	371	SpIN triage test
TPUS	988 (11)		88.4 (74.8 to 96.6)	⊕⊕⊕⊕ High[c]	89.1 (80.8 to 95.9)	⊕⊕⊕⊕ High[c]	521	45	68	366	SpIN triage test
EVUS	454 (2)		69.0 (51.5 to 88.8)	⊕⊕⊖⊖ Low[d]	76.5 (53.5 to 92.9)	⊕⊖⊖⊖ Very Low[d]	407	97	182	314	SpIN triage test; quality of evidence too low to recommend use
DAE	99 (2)		74.6 (53.8 to 91.6)	⊕⊖⊖⊖ Very low[e]	88.5 (61.6 to 98.5)	⊕⊖⊖⊖ Very low[e]	568	45	21	366	SpIN triage test; quality of evidence too low to recommend use
EDF	169 (4)		96.4 (86.8 to 99.4)	⊕⊕⊖⊖ Low[f]	89.0 (59.7 to 98.7)	⊕⊖⊖⊖ Very low[f]	439	47	150	364	SpIN triage test; quality of evidence too low to recommend use
Enterocele											
EP	2233 (31)	24.1 (19.6 to 28.7)	91.2 (83.2 to 97.1)	⊕⊕⊕⊕ High[g]	96.5 (93.4 to 98.9)	⊕⊕⊕⊕ High[g]	220	27	21	732	N/A
MRI	1222 (17)		84.5 (71.8 to 94.0)	⊕⊕⊕⊖ Moderate[h]	99.2 (96.3 to 99.9)	⊕⊕⊕⊕ High[h]	204	6	37	753	SpIN triage test
TPUS	976 (10)		83.6	⊕⊕⊕⊖	98.4	⊕⊕⊕⊕	201	12	40	747	SpIN triage test

			(63.1 to 96.0)	Moderate[i]	(95.1 to 99.8)	High[i]					
EVUS	471 (3)		67.7 (51.2 to 91.4)	⊕⊕⊖⊖ Low[j]	96.9 (80.2 to 99.2)	⊕⊕⊕⊖ Moderate[j]	163	24	78	735	SpIN triage test
DAE	70 (2)		74.5 (52.4 to 94.3)	⊕⊕⊖⊖ Low[k]	96.8 (75.2 to 99.6)	⊕⊕⊕⊖ Moderate[k]	171	20	70	739	SpIN triage test
EDF	139 (3)		70.9 (51.2 to 95.9)	⊕⊖⊖⊖ Very low[l]	97.4 (86.9 to 99.6)	⊕⊕⊖⊖ Low[l]	179	24	62	735	SpIN triage test; quality of evidence too low to recommend use.
Intussusception											
EP	1613 (27)	44.1 (34.7 to 52.6)	88.8 (78.8 to 96.3)	⊕⊕⊕⊕ High[m]	91.8 (85.9 to 97.2)	⊕⊕⊕⊕ High[m]	392	46	49	513	N/A
MRI	480 (11)		60.6 (50.8 to 78.1)	⊕⊕⊕⊕ High[n]	96.7 (88.1 to 99.5)	⊕⊕⊕⊕ High[n]	267	18	174	541	SpIN triage test
TPUS	664 (10)		75.0 (53.6 to 92.8)	⊕⊕⊕⊖ Moderate[o]	96.4 (90.9 to 99.1)	⊕⊕⊕⊕ High[o]	331	20	110	539	SpIN triage test
EVUS	454 (2)		63.2 (51.1 to 87.5)	⊕⊕⊖⊖ Low[p]	92.6 (71.5 to 98.7)	⊕⊕⊕⊖ Moderate[p]	279	41	162	518	SpIN triage test
DAE	99 (2)		61.4 (50.5 to 89.2)	⊕⊖⊖⊖ Very low[q]	92.7 (64.6 to 99.0)	⊕⊕⊖⊖ Very low[q]	271	41	170	518	SpIN triage test; quality of evidence too low to recommend use
EDF	169 (4)		89.3 (65.1 to 98.5)	⊕⊖⊖⊖ Very low[r]	92.4 (71.9 to 98.9)	⊕⊖⊖⊖ Low[r]	394	43	47	516	Replacement test; quality of evidence too low to recommend use
Anismus											
EP	985 (15)	24.8 (18.5 to 31.6)	80.4 (63.1 to 93.7)	⊕⊕⊖⊖ Low[s]	96.8 (94.4 to 98.8)	⊕⊕⊕⊕ High[s]	199	24	49	728	N/A

MRI	287 (7)		85.9 (60.4 to 98.2)	⊕⊖⊖⊖ Very low[t]	95.8 (89.4 to 98.6)	⊕⊕⊕⊖ Moderate[t]	213	32	35	720	SnOUT triage test; quality of evidence to low to recommend use
TPUS	651 (5)		91.9 (72.1 to 98.3)	⊕⊕⊕⊕ High[u]	91.3 (83.1 to 96.7)	⊕⊕⊕⊕ High[u]	228	66	20	686	SnOUT triage test
EVUS	454 (2)		84.5 (59.1 to 96.2)	⊕⊕⊖⊖ Low[v]	90.5 (63.0 to 97.6)	⊕⊕⊖⊖ Low[v]	209	72	39	680	SnOUT triage test; quality of evidence too low to recommend use
DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EFD	169 (4)		87.3 (71.6 to 96.2)	⊕⊕⊖⊖ Low[w]	92.9 (73.8 to 99.1)	⊕⊕⊖⊖ Low[w]	216	54	32	698	SnOUT triage test; quality of evidence too low to recommend use
PFD											
EP	476 (10)	66.9 (55.0 to 78.1)	97.5 (92.6 to 99.5)	⊕⊕⊕⊕ High[x]	83.3 (58.7 to 96.2)	⊕⊕⊕⊖ Moderate[x]	652	55	16	277	N/A
MRI	350 (7)		93.8 (81.4 to 98.4)	⊕⊕⊕⊖ Moderate[y]	79.2 (53.7 to 96.7)	⊕⊖⊖⊖ Very low[y]	627	69	41	263	SpIN triage test; quality of evidence too low to recommend use
TPUS	54 (1)		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
EVUS	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	99 (2)		92.9 (64.4 to 99.1)	⊕⊖⊖⊖ Very low[z]	74.2 (53.6 to 93.4)	⊕⊖⊖⊖ Very low[z]	564	25	104	307	None
EDF	29 (1)		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Risk of bias (QUADAS-2)[8]	Participant selection 'selection bias': High risk: 7 studies; unclear risk: 8 studies; low risk: 24 studies Index test (MRI or Ultrasound) 'interpretation bias': High risk: 3 studies; unclear risk: 8 studies; low risk: 29 studies Index test (EP) 'interpretation bias': High risk: 1 study; unclear risk: 10 studies; low risk: 28 studies Flow and timing 'selection bias': High risk: 5 studies; unclear risk: 7 studies; low risk: 27 studies										

Applicability (QUADAS-2)	Participant selection: Data applicable to women with symptoms of ODS or general pelvic floor dysfunction, or both, presenting to secondary or tertiary care; and to all women regardless of age, parity, body mass index and previous surgery. Data not applicable to asymptomatic women, women presenting to primary care nor to male patients. Index test: Data applicable to different methods of performance of techniques, different cut-off values and level of experience of operators.
Heterogeneity	Sensitivity of MRI performed with an evacuation phase was higher than without evacuation phase, and sensitivity of MRI without evacuation phase was significantly lower than EP for rectocele and enterocele; therefore MRI should be performed with an evacuation phase. Sensitivity of TPUS performed with rectal contrast was not significantly higher than without rectal contrast for rectocele, enterocele and intussusception, and was lower than sensitivities of EP; so it is not recommended for clinical use as it is an invasive procedure and EP remains superior.
Sensitivity analysis	Estimates of sensitivity and specificity calculated in the subset analysis without studies that could reduce overall quality of the evidence (e.g. excluding studies with high risk of bias, concerns about applicability and studies published before 2010) were not notably different compared to the main analysis. Overall, based on all diagnostic tests and target conditions, the median difference was for sensitivity −2.6% (IQR −7.1% to −0.5%) and for specificity −1.5% (IQR −3.9% to −0.2%).
Conclusion	EP remains the best diagnostic imaging technique and cannot be replaced. MRI and TPUS could be used as a triage test, as a positive test confirms a diagnosis of rectocele, enterocele and intussusception, and a negative test rules out diagnosis of anismus. Quality of evidence of EVUS, DAE and EDF was too low to support recommendations.

CrI = Credibility interval; N/A is not analysable

[1]The imaging techniques must be used in a population of women with symptoms of obstructed defaecation syndrome, i.e. difficulty in evacuating stools from the rectum, the sensation of incomplete emptying or the need to digitate to empty, or both. We exclude studies in men and asymptomatic women.

[2]Studies must include EP and any other index test(s).

[3]Although we include studies that used EP as reference standard, EP was taken as an index test similar to the other index tests in the meta-analysis.

[4]Latent class analysis is a modelling technique that allows us to estimate the sensitivity and specificity of a set of diagnostic tests in situations in which there is no good reference standard.

[5]The numbers of EP are based on the a sum of all comparisons of index tests to EP (MRI, TPUS, EVUS, EDF and DAE). Some studies examined more than one index test.

[6]GRADE quality of the evidence:

High: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

[7]Whether or not it meets criteria of replacement test, SpIN triage test or SnOUT triage test with clinical implication.

[8]QUADAS-2 is a tool for the assessment of methodologic quality. The tool comprises four domains: participant selection, index test, reference standard, and flow and timing. In this review the domain 'reference standard' has been changed to 'index test: EP', in the absence of a reference standard. Each domain is assessed for risk of bias, and the first two domains are also assessed for concerns about applicability.

GRADE assessment footnotes:

[a]No downgrading in any of the four domains for sensitivity. Specificity downgraded by one level for unexplained heterogeneity.

[b]No downgrading in any of the four domains for sensitivity or specificity. Borderline judgement for consistency of specificity.

[c]No downgrading in any of the four domains for sensitivity and specificity. Borderline judgement for directness of sensitivity; decrease of 9% in a selected group of women with ODS only.

[d]Sensitivity and specificity downgraded by two levels for unexplained heterogeneity and precision (low number of studies). Also wide credibility intervals were present, but already downgraded for inconsistency. Specificity downgraded by an extra level for directness: specificity increased with 12% in a selected group of women with ODS only. Borderline judgement for directness of sensitivity; decrease of 9%.

[e]Sensitivity and specificity downgraded by three levels: two for precision; small sample size and wide CrI, and one for inconsistency as all studies were from the same unit, although no heterogeneity was present.

[f]Sensitivity and specificity downgraded by two levels for precision (low number of participants) and inconsistency, as all studies were from the same unit, although no heterogeneity was present. Specificity was downgraded by an extra level for precision because of wide CrI and not yet downgraded for heterogeneity. Borderline judgement for ROB of specificity; specificity decreased 8% in the analysis without studies with high ROB.

[g]No downgrading in any of the four domains for sensitivity or specificity.

[h]Sensitivity downgraded by one level for serious effect of ROB; Sensitivity of MRI for enterocele decreased by 11% in the analysis without studies with high ROB. No downgrading in any of the four domains for specificity.

[i]Sensitivity downgraded by one level because of wide CrI in combination with a borderline judgement for heterogeneity. No downgrading in any of the four domains for specificity.

[j]Sensitivity and specificity downgraded by one level for imprecision (low number of studies). Sensitivity was downgraded by an extra level for imprecision because of wide CrI in combination with borderline judgement for heterogeneity. Borderline judgement for directness; sensitivity of enterocele on EVUS increased by 9% in a selected group of women with ODS only.

[k]Sensitivity and specificity downgraded by one level for imprecision for small sample size. Sensitivity was downgraded by an extra level for imprecision because of wide CrI. Borderline judgement for ROB: sensitivity of DAE decreased 9% in the analysis excluding high ROB studies.

[l]Sensitivity and specificity downgraded by two levels for precision (low number of participants) and inconsistency as all studies were from the same unit, although no heterogeneity was present. Sensitivity was downgraded by an extra level for precision because of wide CrI and not yet downgraded for heterogeneity.

[m]No downgrading in any of the four domains for sensitivity or specificity. Borderline judgement for heterogeneity of sensitivity.

[n]No downgrading in any of the four domains for sensitivity or specificity. Borderline judgement for heterogeneity of sensitivity.

[o]Sensitivity downgraded by one level because of unexplained heterogeneity (inconsistency). Wide CrIs were present but already downgraded for inconsistency.

[p]Sensitivity and specificity downgraded by one level for imprecision (small number of studies). Sensitivity downgraded an extra level for precision because of wide CrIs.

[q]Sensitivity and specificity both downgraded by one level for inconsistency (all from same unit) and by one level for imprecision (small number of studies). Senstivity was downgraded by an extra level for imprecision because of wide CI and not yet downgraded for heterogeneity.

[r]Sensitivity and specificity downgraded by two levels for inconsistency; because of unexplained heterogeneity and all from same unit. Both sensitivity and specificity downgraded by one level for imprecision (small number of studies). Wide CrI but already downgraded for heterogeneity.

[s]Sensitivity downgraded by two levels for serious effect of ROB (sensitivity of EP decreased with 17% when excluding high risk of bias studies) and wide CrI. Borderline judgement for directness: Sensitivity of anismus increased with 9% in a selected group of women with ODS only. No downgrading for specificity in any of the four domains.

[t]Senstivity downgraded by three levels: one level for serious effect of ROB (sensitivity of MRI decreased with 10% when excluding high risk of bias studies), and two levels for imprecision (small sample size and wide CrI). Specificity was downgraded by one level for imprecision because of small sample size.

[u]No downgrading in any of the four domains for sensitivity or specificity.

[v]Both sensitivity and specificity downgraded by two levels for imprecision because of small number of studies and wide CrI.

[w]Both sensitivity and specificity downgraded by two levels; one level for inconsistency (all studies from the same unit) and one level for imprecision (small sample size).

[x]No downgrading in any of the four domains for sensitivity. Downgrading specificity by one level for inconsistency (unexplained heterogeneity).

[y]Specificity downgraded by one level because of serious effect of ROB (specificity increased 11% when excluding studies with high risk of bias), and one level for inconsistency (unexplained heterogeneity). Both sensitivity and specificity downgraded by one level for imprecision because of small numbers. For specificity also wide CrI, but already downgraded for unexplained heterogeneity.

[z]Sensitivity and specificity both downgraded by three levels; one for inconsistency as all studies were from the same unit, although no heterogeneity was present, and by two levels for precision (small sample size and wide CrI).

BACKGROUND

Obstructed defaecation syndrome (ODS) is difficulty in evacuating stools, requiring straining at defaecation associated with lumpy or hard stools, having the sensation of incomplete evacuation, a feeling of anorectal blockage/obstruction or the need to manually assist defaecation (Sultan 2017). This could lead to excessive straining, dependence on laxatives or enemas or both, unsuccessful attempts with prolonged periods spent in the toilet, and return visits to it (Santoro 2006). In addition, these people tend to digitally assist evacuation of stool by any of the following techniques: a) rectal digitation (manual extraction of stool from the rectum); b) vaginal digitation (supporting the posterior vaginal wall); or c) splinting (manually supporting the perineum or buttocks during straining, usually with thumb or fingers) (Sultan 2017). These symptoms have a significant effect on social, physical, emotional and sexual well-being, all of which impact on quality of life (Bove 2012; Irvine 2002). This syndrome is also known as anorectal outlet obstruction, evacuatory dysfunction, outlet constipation and pelvic outlet obstruction.

ODS is a sub-category of constipation. The prevalence of constipation in the general population is 27% to 30% when self-reported (Garrigues 2004; Irvine 2002; Pare 2001) and 13% when based on the Rome III criteria for constipation (Papatheodoridis 2010). Constipation contributes to cost for both the patient and society in terms of medications, aids purchased, and loss of work days (Dennison 2005). Constipation can be split into three categories: normal-transit constipation, slow-transit constipation and ODS, which can co-exist in the same person (Lembo 2003). ODS can be distinguished from slow-transit constipation by bowel frequency and stool consistency. Slow-transit constipation is defined as infrequent bowel movements (less than twice a week) and hard stools, whilst people with ODS have at least one defaecation or attempted defaecation a day, and symptoms can also be present with soft stools (Altomare 2008). The severity of ODS symptoms can be assessed with disease-specific validated questionnaires (Altomare 2008; Renzi 2013).

Symptoms of ODS often arise between the ages of 40 and 50, when progressive weakening of the supportive tissue occurs (D'Hoore 2003). ODS is observed in up to half of those with chronic constipation, and is five times more common in women than in men (Noelting 2016). The prevalence of ODS is dependent on the definition used. When defined as 'at least weekly' symptoms of difficulty in bowel movements or digitation or both, the prevalence in the general female population is 20%. In women with symptoms of pelvic floor dysfunction the prevalence of ODS is 32% (Whitcomb 2009). Other studies also found a varying rate of digital assistance for defaecation from 7% in women who seek routine gynaecological care to 38% in women with a stage II or higher posterior vaginal wall prolapse (Kahn 2005; Tan 2005). The prevalence of symptoms of ODS is high in the urogynaecological population; 62% in the total urogynaecological population and 71% in women with stage II or higher posterior vaginal wall prolapse (Guzman Rojas 2016).

The aetiology of ODS can be either functional or mechanical (anatomical). Functional causes could be due to inefficient inhibition of the internal anal sphincter (e.g. Hirschprung's disease, Chagas' disease and hereditary internal sphincter myopathy) or inefficient relaxation of the striated pelvic floor muscles (e.g. anismus, spinal cord lesions and multiple sclerosis). Mechanical causes of ODS can be positioned within the rectum (e.g. rectal tumour, rectocele, rectal intussusception and rectal prolapse) or outside the rectum (pelvic floor descent, enterocele/sigmoidocele, uterine or vaginal prolapse or both) (D'Hoore 2003; Santoro 2006).

The most common mechanical (rectocele, enterocele, intussusception and pelvic floor descent) and functional causes (anismus) affecting the posterior part of the pelvic floor, leading to symptoms of ODS, are called 'posterior pelvic floor disorders'. These can all be visualised with radiologic imaging. As the cause of the symptoms in people with ODS is not always visible on physical examination (Kelvin 1992; Kelvin 1999), it is recommended to perform diagnostic imaging to confirm the diagnosis when posterior pelvic floor disorders are suspected, particularly when considering surgical options (Berman 2005). Inaccurate diagnosis could lead to inappropriate and ineffective surgery with increased risk of complications. Depending on the severity of symptoms and imaging findings, about 20% to 42% of people with ODS need surgical repair, mainly comprising transvaginal rectocele repair, abdominal sacrocolpopexy, ventral rectopexy or stapled transanal rectal resection (STARR) (Kapoor 2008; Podzemny 2015). It has been shown that operative repair reduces ODS symptoms and improves the quality of life (Altomare 2018; Bock 2013; Racaniello 2015; Renzi 2013).

Currently no reliable reference standard exists for the diagnosis of posterior pelvic floor disorders. Evacuation proctography (EP) is the first established and most commonly used diagnostic imaging technique for the assessment of posterior pelvic floor disorders, and is therefore considered to be the reference standard investigation (Sultan 2017); however, the technique has been criticised because of its significant intra- and inter-observer variability (Goei 1989; Müller-Lissner 1998; Van Iersel 2017). Moreover, it has been shown not to deliver perfect test accuracy. When using intraoperative results as the reference standard, sensitivity of EP for rectocele was 50% and specificity was 93%. For enterocele, sensitivity was 47% and specificity 79%, this low test accuracy was caused by enteroceles either missed by EP or wrongly identified as being an enterocele, when it was actually a uterine prolapse (Lienemann 1997). Faucheron 2014 found a sensitivity of 83% and a specificity of 100% for the diagnosis of enterocele and a sensitivity of 88% and a specificity of 100% for diagnosis of intussusception, similar to the test accuracy of MRI when compared to intra-operative findings. Intra-operative findings, however, are not always available as not all patients require surgical repair; moreover, during an operation maximum Valsalva (during which diagnosis is made clinically) is not possible under anaesthesia, and therefore intra-operative findings can not be a potential reference standard. In the absence of a reference standard and when no a priori consensus exists about what combination of tests would be a suitable reference standard, the method to evaluate the accuracy of multiple diagnostic tests is latent class analysis (Rutjes 2007). Latent class analysis combines the results of the imaging techniques through a statistical model to identify the true patient status. It assumes that the actual results of the techniques are imperfect observations of the true unobserved patient status: latent classes 'healthy' and 'diseased'.

Target condition being diagnosed

The target conditions being assessed in this review are the posterior pelvic floor disorders which are visible on radiologic imaging.

Rectocele

The rectovaginal septum is the supportive tissue between the rectum and the vagina (DeLancey 1999; Ludwikowski 2002; Richardson 1993; Zbar 2003), although some authors were unable to find histological evidence of a distinct fascial layer between the anterior wall of the rectum and the posterior vaginal wall (Kleeman 2005). A defect in the rectovaginal septum could allow the anterior rectal wall to herniate through this defect into the vagina. A rectocele is defined as an outward bulge of the rectal wall. This often causes a bulge of the posterior vaginal wall into the lumen of the vagina, which is visible on physical examination. A posterior rectocele (a bulge of the posterior wall of the rectum) is less common and not visible on physical examination. Although childbirth appears to be associated with an increase in the prevalence and size of these defects (Dietz 2006), rectoceles can occur in women who have not been pregnant (Dietz 2005b). A rectocele of less than 2 cm is often asymptomatic and therefore clinically irrelevant (Freimanis 1991; Palit 2014; Shorvon 1989). The prevalence of a rectocele of more than 2 cm in women with symptoms of ODS is 37% to 42% (Martellucci 2011; Weemhoff 2013). A rectocele is clinically significant if it fills preferentially or if it fails to empty after simulated defaecation (Lembo 2003), with a prevalence of 28% in women with defaecatory dysfunction (Hainsworth 2016).

Enterocele

In normal circumstances, the most inferior point of the abdominal cavity is the pouch of Douglas, situated above the vagina and the rectum. The peritoneal sac with abdominal content could herniate between the rectum and the vagina and potentially even protrude into the vagina. Usually, the herniated peritoneal sac contains small bowel loops (enterocele), but sometimes it is filled with the sigmoid colon (sigmoidocele). Enteroceles are divided into posterior, lateral and anterior, depending on which aspect of the vaginal wall is affected (Nichols 1972). The posterior enterocele is by far the most common, with the other two being uncommon (Cronje 2004). There are several factors that contribute to the formation of an enterocele: frequent and prolonged straining, chronically increased intra-abdominal pressure, postmenopausal status, multiparity and previous hysterectomy (Chou 2000; Cronje 2004; Karasick 1997; Lapalus 2004; Mellgren 1994b; Nichols 1972; Oom 2009). The exact prevalence of this condition is unclear. On EP, enteroceles were found in 10% of healthy asymptomatic female volunteers (Shorvon 1989), in 11% to 25% of women with symptoms of pelvic floor dysfunction (Lapalus 2004; Takahashi 2006), and in 23% of women who are investigated for ODS (Morandi 2010). In women who had surgery for pelvic floor disorders, an enterocele was identified in 25% to 45% (Chou 2000; Cronje 2004). The role of enterocele as a causative factor in ODS is controversial, as people with ODS often have a combination of various abnormalities (Morandi 2010).

Intussusception

Intussusception is defined as invagination of the rectal wall into the rectal lumen during defaecation (Dvorkin 2004). It may be described as anterior, posterior or circumferential. The intussusception may involve the full thickness of the rectal wall or only the mucosa. It can be classified as intra-rectal (remains within the rectum), intra-anal (extends into the anal canal), or external (complete rectal prolapse) (Santoro 2011). There is often no identifiable cause in adults, although it appears to be more common in multiparous women, suggesting that it may be associated with other pelvic floor damage (Santoro 2011). Intussusception is the least common cause of ODS, with a prevalence of 4% in the urogynaecological population (Rodrigo 2011). Small intrarectal intussusceptions may be detected in asymptomatic people (Freimanis 1991; Palit 2014; Shorvon 1989), but when the infolding becomes intra-anal, the patient experiences a sensation of incomplete defaecation due to outlet obstruction. In 76% of the women, intussusception is associated with concomitant posterior compartment disorder such as a rectocele, enterocele or anismus (Karlbom 1999; Stoker 2000).

Anismus

In normal circumstances, the pelvic floor and anal sphincter muscles relax during defaecation. Anismus is a state of paradoxical pelvic floor contraction during attempts to evacuate, resulting in inadequate rectal emptying. Previously, other terminology has been used to describe this abnormality: pelvic floor dyssynergy, paradoxical puborectalis syndrome, spastic pelvic floor syndrome and non-relaxing puborectalis syndrome (Stoker 2000). Its exact prevalence is unknown (D'Hoore 2003). Anxiety and psychological stress may contribute to the development of anismus and it is more common in women with a history of sexual abuse (Leroi 1995). In contrast to the other posterior pelvic floor disorders, the incidence of anismus decreases with age (Murad-Regadas 2012a).

Pelvic floor descent

The descending perineum syndrome was first described by Parks 1966, and is associated with abnormal descent of the perineum on clinical examination. It is caused by weakening of the pelvic floor muscle as the result of either neuropathic degeneration, trauma during pregnancy/childbirth or permanent damage due to exaggerated defaecation efforts (Barthet 2000). It is a complex pelvic floor disorder, usually associated with pelvic organ prolapse. Excessive perineal descent may cause stretch injury to the pudendal nerves and sacral roots (denervation) leading to the development of a neuropathy-related faecal incontinence (Bartolo 1983). The synonym of perineal descent on imaging is called pelvic floor descent and is defined as abnormal descent of the ano-rectal junction during straining.

Index test(s)

Evacuation proctography

Evacuation proctography (EP) enables dynamic evaluation of the anatomy and function of the anorectum and pelvic floor during defaecation. Conventionally, EP is performed in the sitting position using barium paste as rectal contrast. The defaecation process is evaluated during evacuation of the contrast using X-ray. EP is simple to perform and widely available, but it involves exposure to ionising radiation, which should be avoided in women who are or might become pregnant, because of its risk of teratogenicity (Williams 2010). EP requires preparation of the small bowel with oral contrast and the large bowel with rectal contrast. It is a lengthy procedure, as the small bowel preparation should be given to the patient one hour prior to the actual investigation. This preparation could cause nausea and constipation as a side effect. The outcome of this investigation is dependent on the consistency of the rectal contrast, which may vary depending on the time between preparation and usage, and the patient's effort to evacuate the contrast. Most women will find the bowel preparation

and the need to defaecate in a non-private setting embarrassing and unpleasant.

Dynamic magnetic resonance imaging

Magnetic resonance imaging (MRI) is an investigation which does not involve ionising radiation. Dynamic MRI (MR-defaecography) is capable of visualising soft tissue and pelvic muscles in different planes and is therefore highly sensitive to detect anatomic abnormalities of the anterior (bladder and urethra), middle (vagina and uterus) and posterior (bowel and anorectum) compartment, whereas with EP (without opacification of the bladder or vagina or both) little or no information about the anterior and middle compartment can be obtained (Stoker 2001). Like EP, MRI is a dynamic investigation and allows assessment of the rectum at rest, during straining and evacuation after the application of rectal contrast (ultrasound gel). Unlike EP, because of excellent tissue discrimination of MRI, no contrast is required for the small bowel, vagina or bladder, which considerably reduces the preparation time. The disadvantages of dynamic MRI are that it is an expensive method, and not widely available. MRI may not be suitable for patients with metal implants (e.g. aneurism clips, pacemakers, sacral nerve stimulation implants, etc.) and with claustrophobia. Unlike EP, the patient is normally in the supine position during examination, which does not mimic the physiological defaecation position. Although open-magnet MR defaecography can be performed with the patient in the sitting position, it requires a specially-designed open scanner, which is even more expensive and restricted to large medical centres (Bertschinger 2002).

Pelvic floor ultrasound

Ultrasound is a non-invasive investigation without use of ionising radiation. Compared to other imaging techniques the costs of ultrasound are relatively low and include a fixed purchase price of the ultrasound scanner and variable overheads for maintenance, ultrasound gel and probe covers. Ultrasound can be an 'in-office' examination and therefore can be done at the same time as the consultation and clinical examination. The benefit of this is that the treating clinician is able to correlate the ultrasound findings with symptoms and clinical findings, but this does require training and expertise. Ultrasound is more readily available and is a faster investigation compared to either EP or MRI. There are no risks or adverse events associated with the use of ultrasound and it is better tolerated by patients as no bowel preparation or contrast is required. Ultrasound provides real-time imaging of the pelvic structures, allowing for static and dynamic investigation of all three compartments. There are various types of ultrasound techniques, which are used for the assessment of the pelvic floor: transperineal ultrasound (TPUS) is performed with a probe placed on the perineum; endovaginal ultrasound (EVUS) with a probe placed into the vagina; and endorectal ultrasound with a probe inside the rectum. The latter is not to be confused with a probe in the anal canal (endoanal ultrasound; EAUS) which is used to assess anal sphincter injuries.

TPUS, using a curved array transducer, enables both 2D and 4D imaging with multiplanar or tomographic reconstructions in any freely-definable plane, and has excellent tissue discrimination (Dietz 2014; Wieczorek 2011). EVUS and endorectal ultrasound are performed with a high-frequency linear probe, which is closer to the area of interest than TPUS. The high resolution results in fluent and representative views of the pelvic floor anatomy (Santoro 2011; Shobeiri 2012). The use of endorectal ultrasound is limited in the assessment of the anterior and middle compartments (Regadas 2011). The disadvantages of ultrasound in general are possible discomfort and embarrassment to the patient due to the position of the probe. The probe and the supine or left-lateral position of the patient may restrict the complete descent of prolapse during straining, thereby underestimating the degree of prolapse. Another limitation of ultrasound is that there is normally no evacuation phase, which is very important, as some abnormalities are more likely to become apparent at the end of evacuation (e.g. full-thickness rectal prolapse and recto-anal intussusception), and incomplete evacuation is not assessed. Moreover, as pelvic floor ultrasound is operator-dependent, it should be performed after an adequate learning curve (Beer-Gabel 2004; Santoro 2011; Wieczorek 2011).

Clinical pathway

People with symptoms of ODS are referred by the general practitioner to a colorectal surgeon or a (uro)gynaecologist, or both. No prior testing is performed. A history is taken to allow the specialist to differentiate between slow-transit constipation, ODS or a combination of both. Women with normal-transit constipation are not usually referred because symptoms may be controlled by laxatives. Specific questionnaires and pathways can help to assess complex symptoms and the impact on the patient's quality of life (Altomare 2008; Renzi 2013; Sultan 2017).

On physical examination, a rectocele can be diagnosed and staged using the standardised POP-Q method (Bump 1996) for the assessment of pelvic organ prolapse. The stage of rectocele can be defined using the following classification: Stage 0 = no rectocele; Stage 1 = most distal part of rectocele is > 1 cm above the level of the hymen; Stage 2 = most distal part of rectocele is between ≤ 1 cm proximal or distal to the plane of the hymen, and Stage 3 = most distal part of rectocele is > 1 cm below the plane of the hymen. Colorectal surgeons usually do not use the POP-Q method and diagnose a rectocele by performing a rectal or a vaginal examination, or both (Beggs 2014).

The diagnosis of an enterocele is based on identifying a hernia sac between the vagina and rectum. This can be achieved by a combined vaginal and rectal digital examination (particularly in the upright position). It is important to differentiate between rectocele and enterocele, as it changes the surgical options, but an enterocele cannot always be detected by physical examination (Kelvin 1999). Rigid sigmoidoscopy should be a part of clinical examination to evaluate intussusception, but this investigation is usually performed by gastroenterologists or colorectal surgeons.

Intussusception is not easily demonstrable clinically, but pre-operative identification is important as it may modify the surgical approach in patients with concomitant posterior pelvic floor disorders (Weemhoff 2013).

Anismus may be identified by digital rectal examination. It is suspected in the absence of relaxation or further tightening of the anal canal during attempts to strain (paradoxical contractility).

Pelvic floor descent is visible on clinical examination when the perineum is ballooning downwards beyond the ischial tuberosities during straining (hence also known as perineal descent).

Additional investigations for the assessment of ODS are the balloon expulsion test, anorectal manometry and electromyography (EMG). Anorectal testing is recommended in patients in whom conservative management in the form of diet, lifestyle modification, empiric laxatives treatment and biofeedback have failed (Bharucha 2014). In patients with ODS who are suspected of co-existing conditions, additional investigations can be requested: bowel transit-time study using X-ray to assess slow-transit constipation, colonoscopy to exclude malignant conditions of the large bowel, endoanal ultrasound and anorectal manometry for the assessment of anal incontinence.

Additional radiological imaging needs to be considered in people with symptoms of difficulty in bowel evacuation or digitation or both, or when posterior pelvic floor disorders are suspected, especially when considering the surgical options (Berman 2005). Currently EP and MRI are used in clinical practice. In some centres, EP has been substituted with dynamic MRI without evidence of similar accuracy. Pelvic floor ultrasound for the posterior pelvic floor is largely used as a research tool, and is emerging into clinical practice.

All possible diagnostic options for the detection of ODS are: validated questionnaires, physical examination, anorectal manometry, balloon expulsion test, electromyography, pelvic floor ultrasound, MRI and EP.

Rationale

EP is the most commonly used imaging technique for the diagnosis of posterior pelvic floor disorders and has been regarded as the reference standard, because of extensive experience and the lack of a perfect reference standard, even though it has been proven not to have perfect accuracy. Moreover, EP is an invasive, embarrassing and unpleasant investigation, which uses ionising radiation and visualises only one compartment. There is an increasing need to find an alternative test for the assessment of ODS that is more acceptable to the patient. Over the years, research has focused on identifying alternative imaging techniques, which would address these disadvantages and may eventually be able to substitute for EP. Although new imaging approaches have been validated and comparative studies of different imaging techniques have been done, the level of agreement between EP, MRI and pelvic floor ultrasound for the diagnosis of posterior pelvic floor disorders varies widely. Due to this lack of consensus about the accuracy of these imaging modalities, it is necessary to conduct a systematic review of the literature.

OBJECTIVES

To determine the diagnostic test accuracy of EP, dynamic MRI and pelvic floor ultrasound for the detection of posterior pelvic floor disorders in women with ODS, using latent class analysis in the absence of a reference standard, and to assess whether MRI or pelvic floor ultrasound could replace EP.

Secondary objectives

To investigate differences in diagnostic test accuracy in relation to the use of rectal contrast, evacuation phase, patient position and use of different cut-off values for the presence of disease, which could influence test outcome.

METHODS

Criteria for considering studies for this review

Types of studies

Diagnostic test accuracy studies and cohort studies (prospective or retrospective) that compare imaging modalities for the detection of posterior pelvic floor disorders were eligible for inclusion. Studies should have used EP in the assessment of test accuracy; either as reference standard or alongside another index test. If test accuracy was not reported, we requested test accuracy data from the study authors; if these data were not available we excluded the studies. We only considered randomised controlled trials if participants were randomised to receive one or other index test and all participants received EP. We excluded case reports and case-control studies, selecting participants with (and without) a specific target condition, as they are likely to overestimate sensitivity and specificity and may potentially cause bias (Whiting 2013).

Participants

We considered studies that recruited women who were suspected of having posterior pelvic floor disorders, e.g. having symptoms of ODS or other symptoms of pelvic floor dysfunction (posterior vaginal wall prolapse, dyschezia, constipation etc.). We excluded studies recruiting only asymptomatic participants. Studies recruiting women with and without symptoms were considered for inclusion, and only data on women with symptoms were included in this meta-analysis. If these data could not be extracted, we approached the study authors to provide test accuracy data only on women with symptoms. We excluded studies that included participants under the age of 18. Studies which included both male and female participants were considered for inclusion and we tried to retrieve test accuracy data only for the women.

Index tests

Studies included in this review assessed test accuracy of two or more imaging modalities that are able to identify posterior pelvic floor disorders. Imaging modalities that are considered as index test are evacuation proctography (EP), dynamic magnetic resonance imaging (MRI) either performed with an open or closed magnet, and various types of dynamic ultrasound; transperineal ultrasound (TPUS), endovaginal ultrasound (EVUS), echodefaecography (EDF) and dynamic anal endosonography (DAE). The index tests considered could be performed with the participant in any preferable position, with or without the use of rectal contrast, and with or without the use of an evacuation phase, as the secondary objective was to find differences in test accuracy when these tests are performed using different methods. For a specific description of how these imaging modalities are performed, see Table 1.

Target conditions

Studies included in this review assessed one or more of the posterior pelvic floor disorders explained in the sections below. This review investigates the accuracy of the index tests to identify each separate condition. We noted that for most target conditions different cut-off values for test positivity were defined for each imaging technique. For this review we chose to use the lowest cut-off for each condition as the cut-off for test positivity. In this way all different cut-offs could be included, as the secondary objective was to investigate differences in diagnostic test accuracy related to

different cut-off values for the presence of disease. For an overview of the existing classifications of the target conditions, see Table 2.

Rectocele

On EP and MRI, a rectocele is diagnosed when there is bulging of the anterior rectal wall. The depth of the rectocele is measured as the maximum depth of the protrusion perpendicular to the expected contour of the anterior rectal wall (Lienemann 1997; Mellgren 1994a). On TPUS a rectocele is diagnosed when a defect in the rectovaginal septum is present, defined as a sharp discontinuity in the anterior anorectal muscularis, resulting in a herniation of rectal contents into the vagina (Dietz 2005a). The method for depth measurement is as described for EP. For EVUS, the diagnosis of a rectocele is similar to that used for TPUS. On EDF a rectocele is diagnosed as a vertical displacement of the lower rectum during defaecatory effort. The depth is measured as the distance between the initial position of the posterior vaginal wall and that at maximal straining (Regadas 2011). On DAE a rectocele is diagnosed if the ventral rectal wall bulges into the vaginal lumen during straining to defaecate (Vitton 2011). For this review, we define test positivity for all imaging techniques as the presence of a rectocele from 0 mm in depth.

Enterocele

On EP, an enterocele is diagnosed when the small bowel or rectosigmoid descends between the rectum and vagina or an enlarged rectovaginal space is visible (Brusciano 2007; Mellgren 1994a). On MRI, an enterocele is defined as descent of the small bowel or sigmoid colon below the pubococcygeal line (PCL) or into the rectovaginal space (Lienemann 1997; Pescatori 2006). The PCL is the connection between the inferior border of the symphysis pubis and the last horizontal coccygeal joint in the sagittal view. On TPUS and EVUS an enterocele is diagnosed when small bowel loops are visible near the rectal vagina septum (Beer-Gabel 2008). On EDF an enterocele is diagnosed when the small bowel is positioned below the ischiococcygeal line (Murad-Regadas 2011) and on DAE, the presence of a herniation of the peritoneal sac in the rectovaginal space during straining is diagnostic (Vitton 2011). For this review, we defined test positivity for all imaging techniques as the presence of any enterocele, either the presence of the small bowel descending between the rectum and vagina, into the rectovaginal space, or below the PCL.

Intussusception

On all imaging techniques intussusception is diagnosed when an invagination of the rectal wall protrudes into the rectal lumen, anal canal or externally during maximal Valsalva (Mellgren 1994a; Beer-Gabel 2004). The Oxford criteria differentiate between high and low rectal or anal intussusception (Collinson 2008). Test positivity for this review for all imaging techniques was defined as the presence of any intussusception; this could either be partial or circumferential intussusception, extending intra-rectally, intra-anally or externally.

Anismus

Whether the puborectalis muscle relaxes or contracts during straining can be measured in all imaging techniques by calculating the difference between the size of the anorectal angle (ARA) at rest and during straining. The ARA is the angle created by the pull of the puborectalis sling at the level of the anorectal junction. A more acute ARA during defaecation indicates a failed release of the puborectal muscle. Some studies reported a wide variation in the value of the ARA both in normal individuals and in people with anismus (Ferrante 1991; Halligan 1995; Shorvon 1989), hence the use of ARA for the diagnosis of anismus is under discussion. Alternatively, anismus can be defined as paradoxical pelvic floor contraction recognised by looking at the cineloops (Hainsworth 2016; Pilkington 2012). Anismus is diagnosed on EP and MRI when a delayed or incomplete expulsion of rectal contrast due to lack of opening of the ARA or anal canal is visible, as a persistent impression of the puborectalis muscle on the posterior rectal wall or as a paradoxical contraction (Kuijpers 1985; Piloni 2013). On ultrasound, anismus is defined when straining is associated with sharpening of the ARA (Martellucci 2011; Murad-Regadas 2008) or when a paradoxical contraction is present (Hainsworth 2016). Test positivity in this review was defined as the presence of anismus either by a decreasing ARA or paradoxical contraction of the puborectalis muscle.

Pelvic floor descent

Pelvic floor descent is determined by measuring the level of the anorectal junction (ARJ) at rest and during straining. Pelvic floor descent is defined as either descent of the ARJ to more than 2 cm below the PCL at rest or descent to more than 3 cm below the PCL on straining (Bartolo 1983), or as a movement of the ARJ of more than 2.5 or more than 3 cm on Valsalva compared to the resting position (Matsuoka 2000; Murad-Regadas 2011). On ultrasound, pelvic floor descent is diagnosed when the distance between the initial and the final position of ARJ during Valsalva is more than 2.5 cm (Murad-Regadas 2011), more than 3.5 cm (Martellucci 2011), or when the puborectalis muscle descends more than 2 cm on straining (Vitton 2011). We define test positivity in this review as the presence of any pelvic floor descent; ARJ more than 0 mm below PCL or a displacement of the ARJ of more than 2.5 cm between rest and Valsalva.

Reference standards

The result of EP is highly dependent on consistency of the contrast, operator experience and evacuatory effort by the person, which could be reduced due to embarrassment. EP is likely to over- or underdiagnose conditions and consequently does not have a perfect test accuracy; it therefore cannot be considered to be a perfect reference standard. Although we included studies that used EP as a reference standard, the meta-analysis was performed using latent class analysis, in the absence of a reference standard, and EP was taken as an index test similar to the other index tests.

Search methods for identification of studies

Electronic searches

We performed a computer-assisted search in MEDLINE and Embase with the OVID interface on 18 December 2019. We designed similarly-structured search strategies using search terms appropriate for each database (see Appendix 1 and Appendix 2 for search strategies run in MEDLINE and Embase). We have not used any search filters (collections of terms aimed at reducing the number needed to screen) as an overall limiter because those published have not proved sensitive enough (Whiting 2011a). We did not apply any language or date restrictions to the electronic searches. We also searched the Cochrane Library's DARE,

CDSR and HTA databases, focusing on systematic reviews and HTAs, and the CENTRAL Register of Controlled Trials (CENTRAL) for primary studies (Appendix 3). We have also searched the Cumulative Index of Nursing and Allied Health (CINAHL) via EBSCO, as this has an appropriate subject focus (Appendix 4). We identified grey literature through the Science Citation Index and Conference Proceedings Citation Index (Appendix 5). One review author developed the search strategy and a search specialist from Cochrane Colorectal Group has approved the official search.

Searching other resources

We performed a handsearch of the references of all included studies to identify eligible studies missed by electronic searches. Through PubMed, we used the included studies to search for additional studies using the 'Related Articles' feature. We used Google Scholar to search for possible eligible studies that have cited one of the included studies. We contacted institutions known to be involved in research of imaging of posterior pelvic floor disorders to collect data of ongoing research. We searched for ongoing trials on www.clinicaltrials.gov and WHO ICTRP (International Clinical Trials Registry Platform) at apps.who.int/trialsearch/, which has data from all approved registers worldwide. In order to reduce publication bias, we attempted to contact researchers involved in studies with possibly relevant but unpublished data.

Data collection and analysis

Selection of studies

Two review authors (IvG and ASt) independently screened titles and abstracts of identified studies, to determine whether they met the inclusion criteria. We obtained a full-text version of each potentially eligible study, identified by the electronic search or by other methods. Two review authors (IvG and ASt) independently evaluated each study for inclusion or exclusion using a study eligibility screening form (Appendix 6), based on prespecified inclusion criteria, resolving disagreements by discussion and in consultation with a third review author (RT). Studies that were published more than once and with overlapping participant data were all selected. We contacted authors with more than one included study, to avoid possible overlap in participant populations; in cases of overlapping participant data, the data of each population were only used once in the meta-analysis. Studies that reported results in men as well as women were all selected, and we contacted authors to provide results on women only. If these data were not available, studies included in the meta-analysis were required to have at least 75% women (an arbitrary cut-off). We included studies in a language other than English that met the inclusion criteria, and we approached the authors of these studies to provide characteristics of the study together with test accuracy data to overcome language restrictions. If these data were not available, we excluded the study.

Data extraction and management

Two review authors (IvG and ASt) independently performed data extraction, including study characteristics, assessment of methodological quality, assessment of possible sources of heterogeneity and test accuracy results from the included studies by using a standardised data extraction form (Appendix 7). We resolved any discrepancies in extracted data by discussion and consultation with a third review author (RT). For all included studies test accuracy data were extracted by participant in numbers of presence or absence of the target condition for each technique, adding up to a total number of participants with a specific test result pattern for each study (Menten 2015). If these data were not available in the published trial reports, we contacted study authors and requested the missing information.

Assessment of methodological quality

We assessed the methodological quality of each study using Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) (Whiting 2011b), as recommended by Cochrane. This tool is made up of four domains: patient selection, index test, reference standard and patient flow, and was tailored with review-specific questions (see Appendix 8). The domain 'Index test' was filled with the results of EP, MRI and ultrasound. The domain 'reference standard' was not filled in the absence of a reference standard, hence this domain was removed from the form. In the 'Characteristics of included studies' and the ROB figures, the 'reference standard' domain was filled with the results of EP as this domain could not be removed from the RevMan software. Because of the design of this review, all included studies performed EP alongside another index test (MRI or ultrasound). Each domain was assessed for risk of bias, and the first two domains were also considered for applicability. We labelled a domain as 'low risk of bias' when all signalling questions were answered with 'yes', as 'unclear risk of bias' when one or more signalling questions were answered as 'unclear' and none with 'no', and as 'high risk of bias' when one or more signalling questions were answered as 'no'. The latter differed from the protocol, as a domain was previously defined as 'high risk of bias' when all signalling questions were answered with 'no'. Two review authors (IvG and ASt) independently assessed the quality of each study, with all disagreements resolved by consultation with a third review author (RT). The influence of risk of bias on the accuracy of index tests was explored in the sensitivity analysis by excluding studies that had at least one domain classified as high risk of bias.

Statistical analysis and data synthesis

As described in the protocol, we initially planned to conduct a frequentist latent class analysis (LCA), using random effects for sensitivity, specificity and prevalence, following Chu 2009. With this approach we aimed to estimate sensitivity and specificity of a pair of two tests, in our study: an index test and EP (the imperfect reference standard), using only the studies in which the index test was applied. Unfortunately, there were not enough studies to estimate in a valid manner by index test and condition the sensitivity and specificity of both the index test and EP.

Our meta-analysis is based on a Bayesian approach to LCA, described by Menten 2015, which allows for the comparison of several diagnostic tests without reference standard. They propose a hierarchical LCA, where the first level consists of a model to describe the observed data per study. The basic data format for this model is not the number of true positives and true negatives for each test T_j, as there is no reference standard, but rather the number of participants that show a certain pattern of outcomes across the J tests performed in a study (pattern-per-participant format). The number of participants with pattern y = (y_1, y_2, . . . , yJ), with y_j the observed binary outcome (0 = negative, 1 = positive) for test T_j, can be denoted as N_y. N_y is assumed to follow a multinomial distribution: $N_y \sim Mult(N, P(y))$, with N the total sample size and P(y) the probability of y, depending of the test accuracy and the prevalence of the target condition. The prevalence of the

target condition is allowed to vary across studies. At the second level a model for the study-specific sensitivity-specificity pairs is specified. Considering the low number of studies with more than two tests (maximum three studies), and after discussing this by email with Menten, we decided to use model 4 from the article. Consequently, the sensitivities and specificities of all diagnostic tests are modelled using separate bivariate normal distributions per test. With this approach we estimated per target condition the pooled sensitivity, specificity, prevalence, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR +), negative likelihood ratio (LR−) and diagnostic odds ratio (DOR) with corresponding 95% credibility intervals (CrIs) of the separate diagnostic tests. PPV, NPV, LR and DOR estimates are based on the estimated pooled prevalence, sensitivity, and specificity. Using the same hierarchical approach, the sensitivity, specificity and prevalence are estimated per study.

We estimated and reported the median parameter values and their 95% CrIs. The median or the 50% quantile is the value below which lies 50% of the posterior sample, which is robust in case some parameters may have skewed posterior distributions. The 95% CrI is the Bayesian equivalent of the classical (frequentist) 95% confidence interval (CI). Under the Bayesian approach, all unknown parameters must be provided a prior probability distribution that defines the range of possible values of the parameter and the likelihood of each of those values based on prior 'beliefs', i.e. information external to the data (Horne 2019). In order to let the observed data determine the final results, we used low-information prior distributions for the pooled sensitivity and specificity parameters and their between-study standard deviation parameters. It is known that Bayesian models can be sensitive to the choice of prior distributions. We therefore carried out additional analyses with alternative prior distributions, e.g. with priors reflecting higher sensitivity and specificity for EP. We noted no appreciable change in pooled accuracy parameters except, as expected, slightly different posterior credible intervals. Parameters of the model were estimated using Markov-Chain Monte-Carlo (MCMC) methods through Gibbs sampling. We summarise the model we used in the Statistical Appendix (Appendix 9).

A new test may replace a current test (EP) when sensitivity and specificity of the new test are similar or higher than those of the current test. To evaluate whether or not the tests under evaluation in this review are able to replace EP for diagnosis of the target conditions, we used the following approach. Based on the MCMC results of the LCA, we calculated the difference between the sensitivity of an index test and EP. We calculated the probability that an index test has an equal or better sensitivity than EP as the proportion of all differences that was ≥ 0. We used the same approach for specificity. A low probability, e.g. 0.10, suggests that the probability that the index test is as least as good as EP is very small. A probability around 0.50 (say 0.40 to 0.60) suggests that the index test performs similarly to EP. A high probability (say 0.90) suggests that the probability that the index test is better than EP is very high. For this review, we defined the accuracy of the index test as similar or higher than EP if the probability was higher than 0.40. An index test is considered suitable as a replacement test if both sensitivity and specificity are estimated to be similar or higher than those of EP (probability more than 0.40 for both).

A SpIN triage test (SpIN: mnemonic to indicate that a highly-specific test (Sp) acts to rule in the condition (in)) has a high specificity, following a low number of false-positives. The high positive predicted value shows that a positive test result confirms the diagnosis with high certainty and that no further testing is needed. However, if no abnormality is found additional testing is necessary to rule out the disease. A SnOUT triage test (SnOUT: Mnemonic to indicate that a highly-sensitive test (Sn) acts to rule out the condition (out)) has a high sensitivity, following a low number of false-negatives. The high negative predicted value shows that a negative test result excludes the disease with high certainty and no further testing is necessary. However if an abnormality is found additional testing is needed to confirm the finding, e.g. rule in the disease (Nisenblat 2016). For this review, we defined an index test as suiting a SpIN triage test if specificity is similar or higher than EP (probability greater than 0.40 for specificity, no restrictions for sensitivity); a positive test result rules in the diagnosis. An index test would suit as a SnOUT triage test if the sensitivity is similar or higher than EP (probability greater than 0.40 for sensitivity, no restrictions for specificity); a negative result rules out the diagnosis.

Visualisation of the data and the results of the main analyses were done with forest plots and summary receiver operating characteristic (ROC) curves. For each index test and condition, we created forest plots based on the results of the LCA approach. In order to use the Review Manager 5 (RevMan) software for the LCA forest plots, we provided for each study the sample size and the median estimates of the prevalence, sensitivity and specificity. RevMan translated these numbers as true-positive, false-positive, true-negative, false-negative numbers and 95% confidence intervals (CIs) for the sensitivity and specificity per study. These CIs were not exactly equal to the 95% CrI produced by the Bayesian analyses.

The ROC curves visualise the LCA results by condition, presenting the pooled result with precision and variation across studies. Bivariate model parameters (E(logitSE), E(logitSp), Var(logitSe), Var(LogitSp), Cov(logits), Corr(logits)) and precision measures to generate confidence regions (SE(E(logitSE)), SE(E(logitSp)), Cov(Es), number of studies) from the Bayesian LCA were used for these ROC curves (Reitsma 2005). The ROC curves were created by RevMan software, apparently using a t-distribution based on the number of studies. The 95% confidence regions in the ROC curves are therefore wider than the CrI as reported from the LCA.

For the MCMC analyses we used OpenBUGS version 3.2.3 (Lunn 2009) in combination with the statistical software R version 3.6.1 (R Core Team 2019), using the packages R2WinBUGS version 2.1.21 (Sturtz 2005), BRugs version 0.9.0 (Thomas 2006), and coda version 0.19.3 (Plummer 2006), using three chains and, depending on the analysis, between 50,000 and 100,000 sampling iterations, with a burning of half the number of iterations. We used RevMan software version 5.4 for reporting and visualisation.

Investigations of heterogeneity

We investigated heterogeneity in the first instance through visual examination of forest plots and ROC graphics of the estimated sensitivities and specificities. To explore the between-study variability, we assessed the following possible sources of heterogeneity: the use of rectal contrast, evacuation phase, participant position, and use of different cut-off values for the presence of disease. If sufficient data were available, we also conducted latent class analyses to assess the effects of the test conditions on the diagnostic test accuracy. Other variables

described in the protocol that could lead to heterogeneity were not assessed, i.e. use of vaginal contrast, type of ultrasound probe/machine, and operator experience.

In the protocol we aimed to assess variation in performance of EP, as it can be performed using bladder, vaginal and/or small bowel contrast to enhance visualisation of the anterior or middle compartments, or both. As this review specifically concerns posterior pelvic floor disorders, the administration of contrast in other compartments, apart from the posterior compartment, was not expected to exert significant influence on test accuracy of posterior pelvic floor disorders, and therefore no heterogeneity analysis was performed for multi-compartmental contrast.

In the secondary objective of the protocol we aimed to estimate test accuracy for each test at prespecified thresholds. However due to insufficient data it was not possible to establish test accuracy for each threshold, but only for combined thresholds. The threshold 'presence' was not to be taken into account in this analysis.

In the analysis we aimed to compare the following test conditions that are possibly related to heterogeneity:

- Use of rectal contrast: present versus absent
- Use of evacuation phase: present versus absent
- Position of participant during the investigation: left-lateral, upright, supine or prone
- Cut-off values used for definition of presence or absence of the target condition:
    - Rectocele: rectocele depth > 0 cm (> 0 and > 1 cm) versus > 2 cm (> 2 and > 3 cm)
    - Enterocele: small bowel in recto-vaginal space versus small bowel below the pubococcygeal line (any > 0 cm)
    - Intussusception: any intussusception (including mucosal and partial) versus full-thickness circumferential
    - Anismus: more acute anorectal angle at straining versus paradoxical pelvic floor contraction
    - Pelvic floor descent: anorectal junction below the pubococcygeal line (any > 0 cm) versus difference in ARJ between rest and Valsalva (any mm)

If a study used two different entities of the test condition in the same population (e.g. unknown part of the population had rectal contrast and the rest of the population not), we excluded this study from the heterogeneity analysis.

If a study provided test accuracy data on a factor that we predefined as a possible source of heterogeneity, we extracted test accuracy data of both test conditions (e.g. test accuracy of index test with and without evacuation phase) in the pattern-per-participant format as required for the LCA, and used them in the heterogeneity analysis. We selected only one of these test conditions for the main analysis, being the one that is most clinically common.

We asked study authors to provide additional test accuracy data on factors that were predefined as a possible source of heterogeneity if available (e.g. DTA for cut-off rectocele > 0 mm and > 20 mm depth). We used the parameter used in the published article for the main analysis and used the additional data, including the original data, for the heterogeneity analysis.

The statistical analyses for heterogeneity were conducted in a similar Bayesian LCA as the main analyses, but with a small variation: to include the various conditions of a diagnostic test in a heterogeneity analysis, we considered the variations of the test as separate diagnostic tests. For example, when evaluating the test accuracy of MRI with and without evacuation phase, we conducted a LCA using pattern-per-participant data for seven tests: EP, MRI with, and MRI without evacuation phase, TPUS, EVUS, DAE and EDF. From this analysis we derived pooled sensitivity and specificity estimates for seven tests. When there were sufficient data to evaluate variations of three tests (i.e. EP, MRI and TPUS), such an analysis was based on models for nine tests (EP 2x, MRI 2x, TPUS 2x, EVUS, DAE and EDF).

Sensitivity analyses

We performed a sensitivity analysis to determine the influence of risk of bias on the diagnostic test accuracy, by excluding studies that were classified as high risk of bias in any of the four domains according to the QUADAS-2 checklist, and by re-analysing the data without these studies.

We performed a second sensitivity analysis to assess the influence on the diagnostic test accuracy of studies included with concerns about applicability. Studies that included participants with symptoms of general pelvic floor dysfunction and studies that included one or more men were excluded from this analysis. If study authors provided data on women only with symptoms of ODS, we included them in the analysis.

We performed a third sensitivity analysis to assess the effect of advancing imaging techniques over time on the diagnostic test accuracy. Ultrasound and MRI techniques develop quickly with improved image quality which enables easier diagnosis of the target conditions. Moreover, different methodology for performance of techniques are studied in the early years to find out how these imaging techniques are best conducted. Recently a more similar protocol for imaging assessment has been used, potentially leading to less heterogeneity. Studies included in this analysis are studies published in the last 10 years (i.e. 2010 and later).

We conducted a fourth sensitivity analysis combining the criteria of the first three sensitivity analyses to establish diagnostic test accuracy excluding studies with high risk of bias, studies with concerns about applicability and by including only studies with recent well-established methodology.

The sensitivity analyses were conducted with the same Bayesian LCA approach as described for the main analyses, conducted in the above explained subsets of studies.

Assessment of reporting bias

We did not assess possible reporting bias. Diagnostic test accuracy reviews show more heterogeneity in included study designs than intervention reviews (mainly RCTs). RCTs are more likely to be published when results are positive or significant, and publication bias could pose an important threat to the validity of these intervention reviews. The subject 'imaging techniques for assessment of ODS' allows for publication of favourable and non-favourable results and it is less likely that reports remain unpublished. There is no evidence of reporting bias in test accuracy reviews and methods to detect this are not very reliable for diagnostic test data (Deeks 2005).

Assessment of overall quality of evidence

We used the GRADE approach for diagnostic test accuracy (Hsu 2011; Singh 2012) to evaluate the overall quality of the body of evidence for each outcome across all studies. We assessed the quality of the evidence using the four factors that may lead to downgrading the quality of the evidence: limitations in study design or execution (risk of bias), directness of the evidence, consistency of the results, and precision. Publication bias was not assessed in this review, and dose-response association, existence of plausible unmeasured confounders and strength of association were not assessed, as these are not applicable for DTA reviews. Observational studies are the next best method after RCTs for quality of the evidence according to the GRADE Working Group (Guyatt 2008). As diagnostic test accuracy is not assessed in RCTs, the default level of evidence is high.

To evaluate overall risk of bias, the results from the QUADAS-2 'Risk of bias' assessment and sensitivity analysis excluding studies at high risk of bias were taken into account. Directness was evaluated by assessing the results from the QUADAS-2 concerns about applicability section and from the sensitivity analysis, excluding studies that caused concerns about applicability. A potential reduction in sensitivity or specificity increases the number of false-positives and false-negatives. As the harm of missed diagnosis and harm of further testing is not high for these benign conditions, since most patients will receive conservative treatment first, we defined a decrease in sensitivity and specificity of more than 10% as being serious, for which we downgraded by one level. Differences close to 10% were defined as borderline and in combination with another borderline judgement the total level of evidence was downgraded by one level. For evaluation of consistency, we explored the forest plots for possible heterogeneity. If heterogeneity was present we checked whether it could be explained by the methods of test performance or methodologic quality. If heterogeneity could be explained no downgrading was necessary, but unexplained heterogeneity caused downgrading of the overall level of evidence by one level for inconsistency. We downgraded by one more level for consistency if the estimated sensitivity or specificity was based on results of studies that were all performed by same research group. For precision, we assessed sample size and credibility intervals. If the analysis was based on three studies or fewer, or contained fewer than 400 participants, we downgraded by one level for imprecision. If credibility intervals of sensitivity or specificity were wider than 30%, we downgraded by one level if the overall level of evidence was not yet downgraded for inconsistency, as this also produces wide credibility intervals. In a consensus meeting with all review authors, we made judgements about the individual criteria and the overall quality of the evidence (high, moderate, low or very low) using a standardised form (Appendix 10).

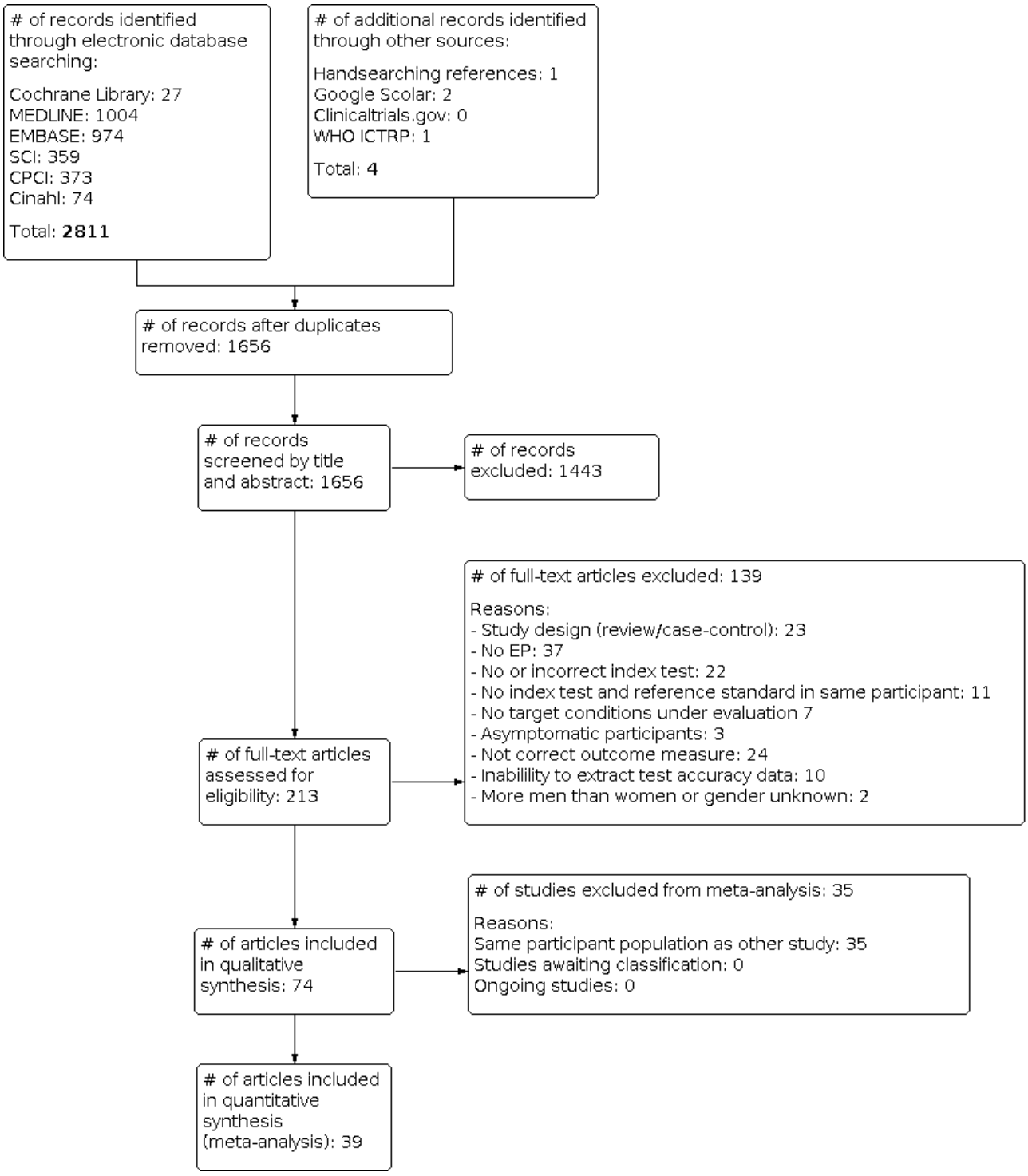
RESULTS

Results of the search

We ran an electronic search on 18 December 2019 in the Cochrane Library, MEDLINE, Embase, SCI, CINAHL and CPCI, resulting in 1656 records after excluding duplicates. Handsearching references, Google scholar and clinicaltrials.gov resulted in four more eligible records. After title- and abstract-screening by two review authors, we retrieved 213 full-text articles and assessed them for eligibility. We excluded 10 studies because of missing test accuracy data and lack of response of the authors to our requests, despite all efforts (Chatoor 2007; Imanova 2017; Kaufman 2001; Kawata 2010; Otto 2011; Pannu 2009; Rizal 2014; Song 2009; Wang 2005; Zeng 2019). One study could be included because authors provided us with test accuracy data (Miravalle 2016). We further excluded two studies because more men than women were included or gender was unknown and test accuracy data on women only were not available from the study authors (Chung 2003; Schoenenberger 1998). For a detailed overview of reasons for exclusion, see Characteristics of excluded studies. Studies were excluded from the meta-analysis if they overlapped in participant population (n = 35, of which 30 were conference proceedings).

In total 39 studies were included in the meta-analysis, with a total of 2483 participants. The study selection process is presented in a PRISMA flow diagram (Figure 1). Most of the included studies were published as full-text articles (37) and two were abstracts of conference proceedings (Ron 2012; Miravalle 2016). In five out of the 39 included studies, all information was available from the published study report. We contacted study authors of the other 34 included studies, to provide additional information to complete the data extraction form. All necessary data were made available by 26 of these authors; four authors were not able to provide the data and we did not receive a reply from another four. A summary of characteristics of the included studies is presented in Table 3.

Figure 1. Study flow diagram (until Dec 18th 2019)



Most of the included studies were prospective single-centre cross-sectional studies. Most studies were performed in Europe (n = 28; 72%) and a few in the Middle-East, North America, South America and Australia. About half of the included studies were published less than 10 years ago (n = 21; 54%). Sample size per study ranged from 7 to 614 participants, with a median sample size of 44. Twenty-one studies investigated MRI as the index test (Table 4) and 20 studies investigated pelvic floor ultrasound as the index test (Table 5); of which 12 investigated TPUS (three from the same unit), three EVUS, three DAE (two from the same unit) and four EDF (all from same unit). Most studies (n = 36; 92%) assessed two tests, and three studies investigated three or more tests. Fewer than half (n = 16; 41%) used EP as the reference standard to calculate sensitivity and specificity of the other imaging technique. The others assessed diagnostic accuracy either without a reference standard (n = 18;

46%), using clinical examination or intra-operative results or both (n = 3; 8%), or a statistical method as reference standard (n = 2; 5%).

For the population being assessed in the included studies, the mean age across studies ranged from 39 to 66 years (median 57.2). Most studies included only women (n = 31; 86%) and of the few studies that included both sexes, in six out of eight, we acquired data for women only. Ethnicity was not reported in most studies (n = 28; 72%). In those that did report ethnicity, most were white (n = 9; 23%), with only two studies reporting mixed ethnicity. Most studies (n = 24; 62%) included participants with symptoms of ODS only; the others included women with a wider range of symptoms including faecal incontinence or pelvic organ prolapse. An overview of definitions of target conditions and cut-off values used in the included studies is presented in Table 6 (rectocele); Table 7 (enterocele); Table 8 (intussusception); Table 9 (anismus); Table 10 (pelvic floor descent).

Methodological quality of included studies

Detailed characteristics and assessment of methodological quality of each individual study are presented in the Characteristics of included studies. The outcome of the methodological quality assessment is presented in Figure 2 and Figure 3, summarising the number of studies with low, high or unclear risk of bias for each of the three domains. Fifteen studies (38%) were classified as low risk of bias in all three domains. Risk of bias was unclear in 12 studies (31%); seven in one domain and five in two/three domains. Twelve studies (31%) were classified as high risk of bias; 10 in one domain and two in more than one domain.

Figure 2. Risk of bias and applicability concerns summary: review authors' judgements about each domain for each included study. The domain 'Index test' could be either Dynamic MRI, Transperineal ultrasound (TPUS), Endovaginal ultrasound (EVUS), Dynamic anal endosonography (DAE) or Echodefaecography (EDF) depending on

what is assessed in the included study. In the domain 'Reference standard' the results for EP are presented. Note that in this review EP is considered as index test and not as reference standard.

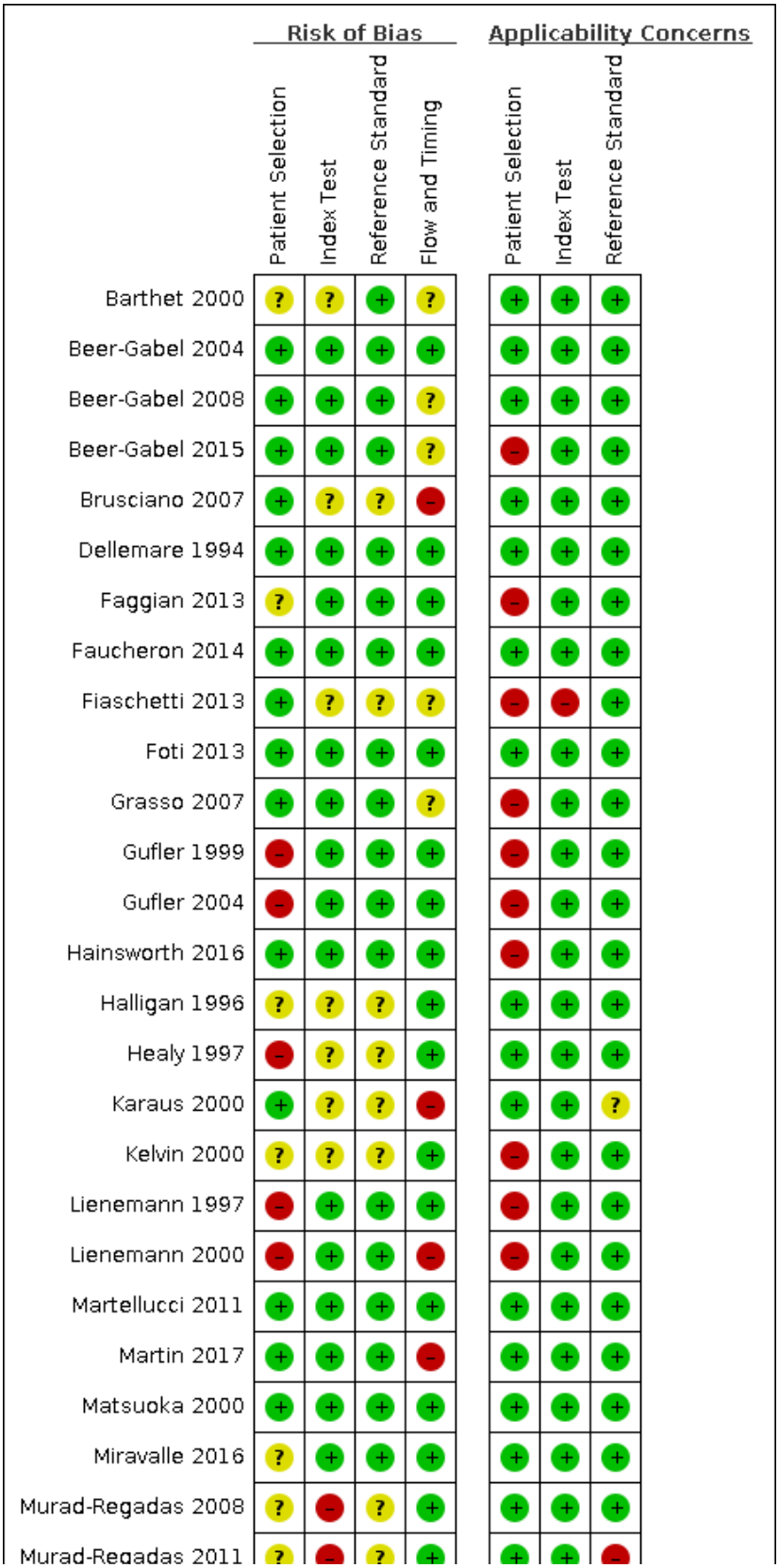


Figure 2. (Continued)

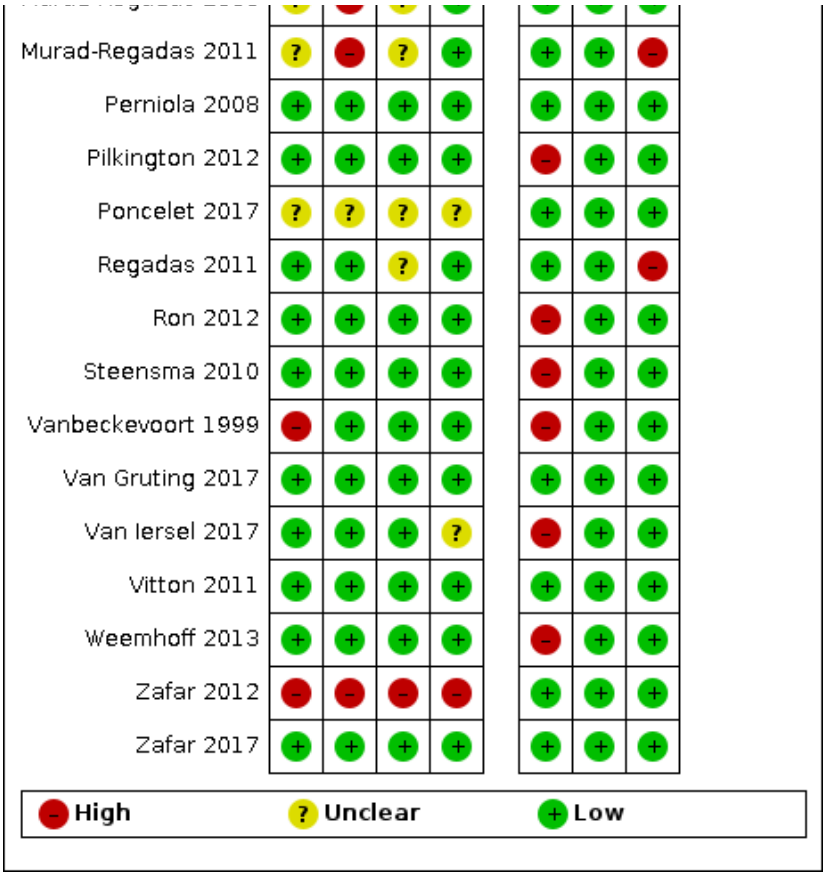
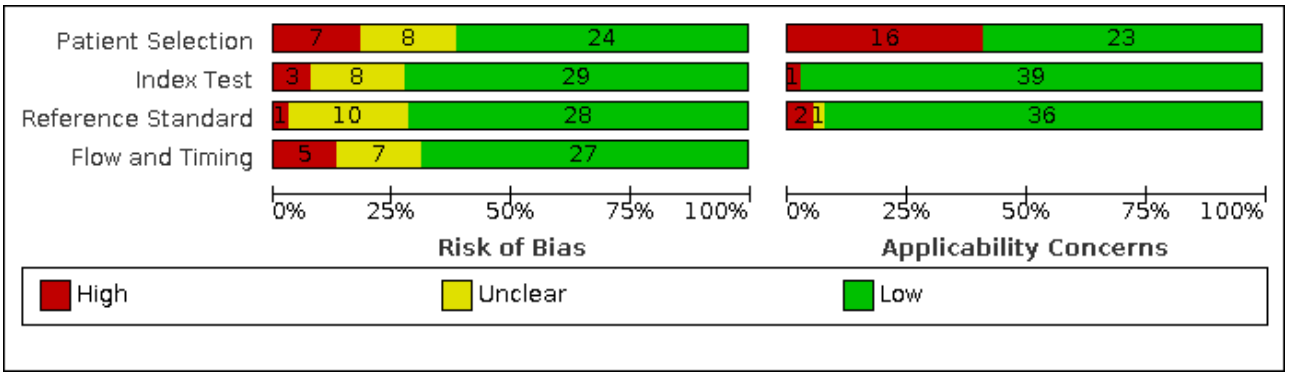


Figure 3. Risk of bias and applicability concerns graph: review authors' judgements about each domain presented as percentages across included studies. The domain 'Index test' could be either Dynamic MRI, Transperineal ultrasound (TPUS), Endovaginal ultrasound (EVUS), Dynamic anal endosonography (DAE) or Echodefaecography (EDF) depending on what is assessed in the included study. In the domain 'reference standard' the results for EP are presented. Note that in this review EP is considered as index test and not as reference standard.



Participant selection

Most of the studies included a consecutive or random sample of women with ODS or general pelvic floor dysfunction. Seven studies were classified at high risk of bias for this domain. Four studies selected participants based on clinical examination (e.g. presence of prolapse or pelvic floor descent), hence increasing the pre-test probability (Gufler 1999; Lienemann 1997; Lienemann 2000; Vanbeckevoort 1999). Two studies selected participants based on having had the imaging technique (retrospective study design), hence excluding women that could have been included (Gufler 2004; Zafar 2012). One study excluded women in whom EP had shown intussusception and rectal prolapse (Healy 1997). In

eight studies it remained unclear if the recruitment process was consecutive or at random, as this was not clearly stated and no exclusion criteria were formulated.

Index test

Most studies analysed the images of the index tests (EP, MRI or ultrasound) blinded, and used pre-established cut-off values. We rated two studies at high risk of bias because cut-off values of echodefaecography for rectocele (Murad-Regadas 2008) and pelvic floor descent (Murad-Regadas 2011) were defined after establishing optimal diagnostic test accuracy. Upon request we received test accuracy data based on any rectocele, and we included these data in the analysis. We rated Zafar 2012 at high risk of bias because analysis of the MRI and EP images was not performed blinded in its retrospective design. Eight studies had an unclear risk of bias for assessment of MRI/ultrasound; in two it remained unclear whether prespecified cut-off values were used (Barthet 2000; Karaus 2000) and in seven it remained unclear whether the analysis was performed blinded (Brusciano 2007; Fiaschetti 2013; Halligan 1996; Healy 1997; Karaus 2000; Kelvin 2000; Poncelet 2017). Ten studies had an unclear risk of bias for assessment of EP; in one it remained unclear whether prespecified cut-off values were used (Barthet 2000) and in ten it remained unclear whether the analysis was performed blinded (Brusciano 2007; Fiaschetti 2013; Halligan 1996; Healy 1997; Karaus 2000; Kelvin 2000; Poncelet 2017; Regadas 2011; Murad-Regadas 2008; Murad-Regadas 2011).

Flow and timing

Most studies included all participants in the analysis that were recruited into the study and used an appropriate timeframe between the tests. Two studies were classified at high risk of bias because we deemed the time interval between tests as inappropriate (mean interval more than three months) (Martin 2017; Zafar 2012) and in eight it remained unclear if there was an appropriate time interval (Barthet 2000; Brusciano 2007; Beer-Gabel 2008; Beer-Gabel 2015; Fiaschetti 2013; Grasso 2007; Poncelet 2017; Van Iersel 2017). We rated three studies at high risk of bias because not all participants were included in the analysis and the exclusions were not explained (Brusciano 2007; Karaus 2000; Lienemann 2000). Some studies did not include all recruited participants in the analysis, but dropouts were explained and were perceived as appropriate/at random, e.g. contra-indication for the index test, withdrawal of patient consent, loss of data sets (Hainsworth 2016; Lienemann 1997; Perniola 2008; Van Gruting 2017; Van Iersel 2017).

Applicability

We present a similar figure for concerns about applicability (Figure 2 and Figure 3). Most concerns are in the section on participant selection. Most studies (n = 24; 62%) included women with symptoms of ODS only, but 15 studies raised a high concern as they did not include participants with only symptoms of ODS, but with a wider range of symptoms including faecal incontinence or pelvic organ prolapse. In clinical practice most patients have a combination of symptoms of pelvic floor dysfunction rather than solely symptoms of ODS, so we decided to include studies that selected women with a wider range of symptoms. We acquired diagnostic test characteristics on women only from six of eight studies that included both men and women. The two other studies raised a high concern about applicability, as they also included men, but the number of included women was above our arbitrary cut-off of 75%, so still met the inclusion criteria. In one study there was a high concern about applicability, as an MRI scanner of only 0.25/0.5 Tesla was used, which could result in poor image quality (Fiaschetti 2013), but this is because it was used upright and higher Tesla is not available. Furthermore, Regadas 2011 and Murad-Regadas 2011 caused concern about the applicability of EP, because the ischiococcygeal line was used as a reference line, whereas all other studies using a reference line used the PCL. Overall, there were some concerns about applicability to the review question, but no concerns about applicability to clinical practice.

Findings

Main analysis

An overview of the diagnostic test characteristics for all imaging techniques can be found in Table 11. We estimated diagnostic test accuracy of all imaging techniques, and we evaluated these tests for their potential to replace EP (replacement test) or to improve selection of women for EP (triage test) that can rule out (SNout) or rule in (SPin) the disease based on probability.

Rectocele

Figure 4 shows the ROC plot and Figure 5 the forest plots for diagnosis of rectocele by all imaging techniques. The estimated pooled sensitivity of EP is 98% (credible interval (CrI) 94% to 99%), of MRI 94% (CrI 86% to 98%), TPUS 88% (CrI 75% to 97%), EVUS 69% (CrI 52% to 89%), DAE 75% (CrI 54% to 92%), and EDF 96% (CrI 87% to 99%) (Table 11). The probability that the sensitivity of the index test is equal to or better than EP is for all tests lower than 0.40 (Table 12). The estimated pooled specificity of EP is 78% (CrI 63% to 90%), of MRI 90% (CrI 78% to 97%), TPUS 89% (CrI 81% to 96%), EVUS 76% (CrI 54% to 93%), DAE 88% (CrI 62% to 98%), and EDF 89% (CrI 60% to 99%). The probability that the specificity of the index test is equal to or better than EP is higher than 0.40 for all. None meet the criteria for replacement test nor for SnOUT triage test. All meet the criteria for SpIN triage test. In case of a positive test, the percentage of women that would truly have a rectocele is 86% for EP, 93% for MRI, 92% for TPUS, 81% for EVUS, 90% for DAE, and 93% for EDF, given a prevalence of 59%. In case of a negative test, 96% for EP, 92% for MRI, 84% for TPUS, 63% for EVUS, 70% for DAE, and 94% for EDF of women will truly not have a rectocele.

Figure 4. Summary ROC Plot from results of the LCA for rectocele for all imaging techniques. Note that the presented confidence regions are wider than the credibility regions reported from the LCA.

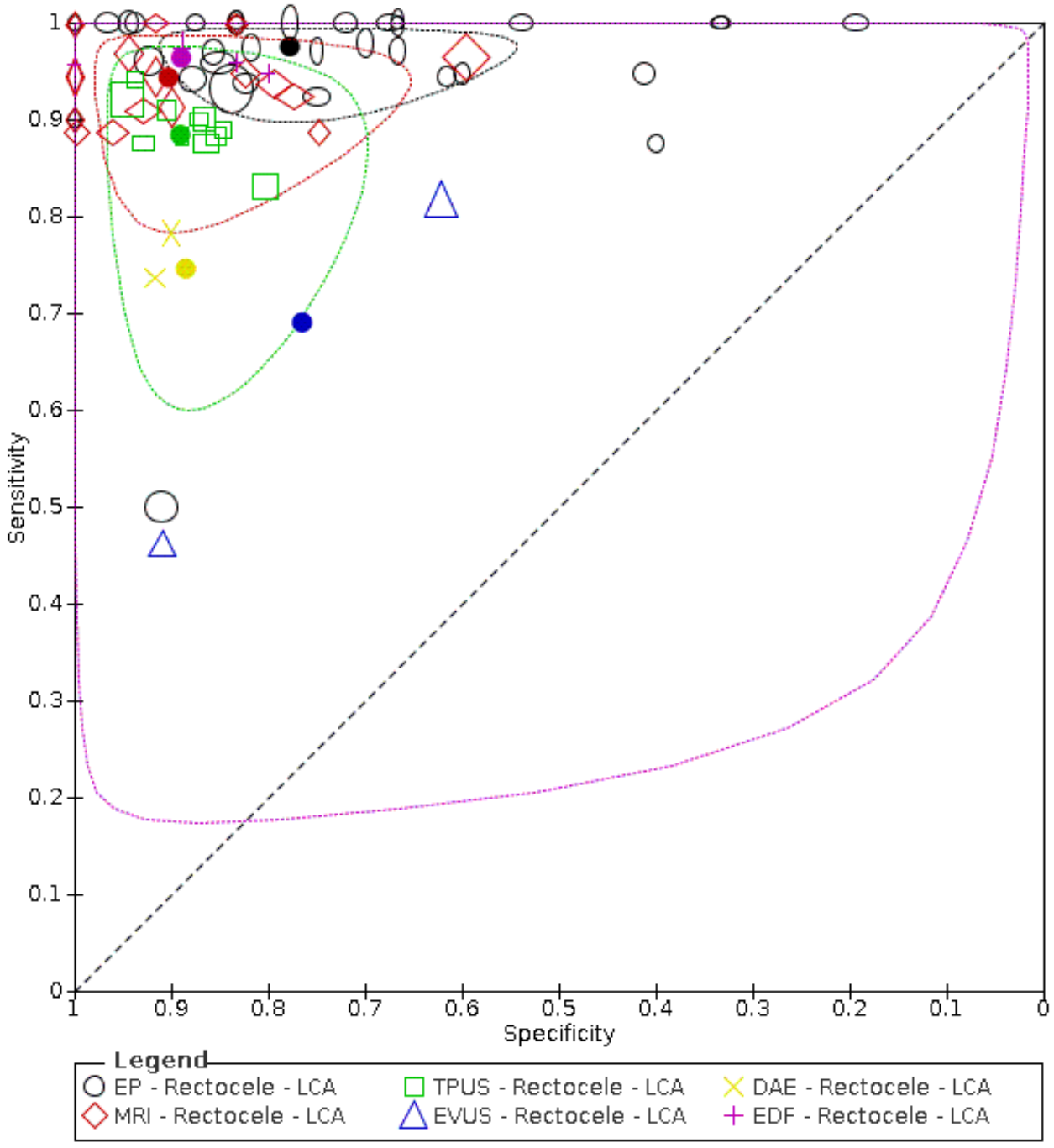
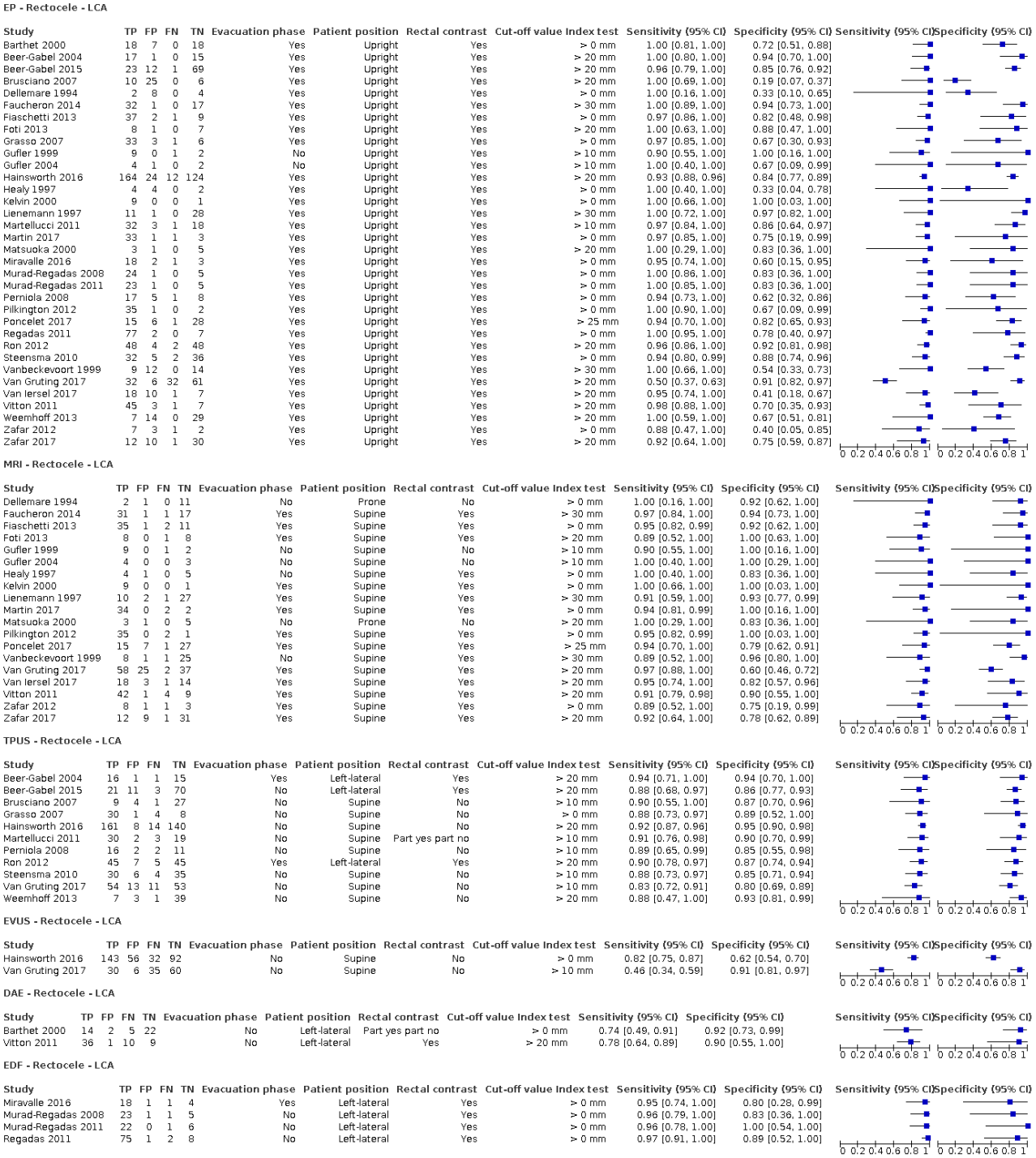


Figure 5. Forest plots of all tests for diagnosis of rectocele based on study specific results of the LCA. Note that these are different from the extracted data.



EP - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	18	7	0	18	Yes	Upright	Yes	> 0 mm	1.00 [0.81, 1.00]	0.72 [0.51, 0.88]
Beer-Gabel 2004	17	1	0	15	Yes	Upright	Yes	> 20 mm	1.00 [0.80, 1.00]	0.94 [0.70, 1.00]
Beer-Gabel 2015	23	12	1	69	Yes	Upright	Yes	> 20 mm	0.96 [0.79, 1.00]	0.85 [0.76, 0.92]
Brusciano 2007	10	25	0	6	Yes	Upright	Yes	> 20 mm	1.00 [0.69, 1.00]	0.19 [0.07, 0.37]
Dellemare 1994	2	8	0	4	Yes	Upright	Yes	> 0 mm	1.00 [0.16, 1.00]	0.33 [0.10, 0.65]
Faucheron 2014	32	1	0	17	Yes	Upright	Yes	> 30 mm	1.00 [0.89, 1.00]	0.94 [0.73, 1.00]
Fiaschetti 2013	37	2	1	9	Yes	Upright	Yes	> 0 mm	0.97 [0.86, 1.00]	0.82 [0.48, 0.98]
Foti 2013	8	1	0	7	Yes	Upright	Yes	> 20 mm	1.00 [0.63, 1.00]	0.88 [0.47, 1.00]
Grasso 2007	33	3	1	6	Yes	Upright	Yes	> 0 mm	0.97 [0.85, 1.00]	0.67 [0.30, 0.93]
Gufler 1999	9	0	1	2	No	Upright	Yes	> 10 mm	0.90 [0.55, 1.00]	1.00 [0.16, 1.00]
Gufler 2004	4	1	0	2	No	Upright	Yes	> 10 mm	1.00 [0.40, 1.00]	0.67 [0.09, 0.99]
Hainsworth 2016	164	24	12	124	Yes	Upright	Yes	> 20 mm	0.93 [0.88, 0.96]	0.84 [0.77, 0.89]
Healy 1997	4	4	0	2	Yes	Upright	Yes	> 0 mm	1.00 [0.40, 1.00]	0.33 [0.04, 0.78]
Kelvin 2000	9	0	0	1	Yes	Upright	Yes	> 0 mm	1.00 [0.66, 1.00]	1.00 [0.03, 1.00]
Lienemann 1997	11	1	0	28	Yes	Upright	Yes	> 30 mm	1.00 [0.72, 1.00]	0.97 [0.82, 1.00]
Martellucci 2011	32	3	1	18	Yes	Upright	Yes	> 10 mm	0.97 [0.84, 1.00]	0.86 [0.64, 0.97]
Martin 2017	33	1	1	3	Yes	Upright	Yes	> 0 mm	0.97 [0.85, 1.00]	0.75 [0.19, 0.99]
Matsuoka 2000	3	1	0	5	Yes	Upright	Yes	> 20 mm	1.00 [0.29, 1.00]	0.83 [0.36, 1.00]
Miravalle 2016	18	2	1	3	Yes	Upright	Yes	> 0 mm	0.95 [0.74, 1.00]	0.60 [0.15, 0.95]
Murad-Regadas 2008	24	1	0	5	Yes	Upright	Yes	> 0 mm	1.00 [0.86, 1.00]	0.83 [0.36, 1.00]
Murad-Regadas 2011	23	1	0	5	Yes	Upright	Yes	> 0 mm	1.00 [0.85, 1.00]	0.83 [0.36, 1.00]
Perniola 2008	17	5	1	8	Yes	Upright	Yes	> 0 mm	0.94 [0.73, 1.00]	0.62 [0.32, 0.86]
Pilkington 2012	35	1	0	2	Yes	Upright	Yes	> 0 mm	1.00 [0.90, 1.00]	0.67 [0.09, 0.99]
Poncelet 2017	15	6	1	28	Yes	Upright	Yes	> 25 mm	0.94 [0.70, 1.00]	0.82 [0.65, 0.93]
Regadas 2011	77	2	0	7	Yes	Upright	Yes	> 0 mm	1.00 [0.95, 1.00]	0.78 [0.40, 0.97]
Ron 2012	48	4	2	48	Yes	Upright	Yes	> 20 mm	0.96 [0.86, 1.00]	0.92 [0.81, 0.98]
Steensma 2010	32	5	2	36	Yes	Upright	Yes	> 0 mm	0.94 [0.80, 0.99]	0.88 [0.74, 0.96]
Vanbeckevoort 1999	9	12	0	14	Yes	Upright	Yes	> 30 mm	1.00 [0.66, 1.00]	0.54 [0.33, 0.73]
Van Gruting 2017	32	6	32	61	Yes	Upright	Yes	> 20 mm	0.50 [0.37, 0.63]	0.91 [0.82, 0.97]
Van Iersel 2017	18	10	1	7	Yes	Upright	Yes	> 20 mm	0.95 [0.74, 1.00]	0.41 [0.18, 0.67]
Vitton 2011	45	3	1	7	Yes	Upright	Yes	> 20 mm	0.98 [0.88, 1.00]	0.70 [0.35, 0.93]
Weemhoff 2013	7	14	0	29	Yes	Upright	Yes	> 20 mm	1.00 [0.59, 1.00]	0.67 [0.51, 0.81]
Zafar 2012	7	3	1	2	Yes	Upright	Yes	> 0 mm	0.88 [0.47, 1.00]	0.40 [0.05, 0.85]
Zafar 2017	12	10	1	30	Yes	Upright	Yes	> 20 mm	0.92 [0.64, 1.00]	0.75 [0.59, 0.87]

MRI - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Dellemare 1994	2	1	0	11	No	Prone	No	> 0 mm	1.00 [0.16, 1.00]	0.92 [0.62, 1.00]
Faucheron 2014	31	1	1	17	Yes	Supine	Yes	> 30 mm	0.97 [0.84, 1.00]	0.94 [0.73, 1.00]
Fiaschetti 2013	35	1	2	11	Yes	Supine	Yes	> 0 mm	0.95 [0.82, 0.99]	0.92 [0.62, 1.00]
Foti 2013	8	0	1	8	Yes	Supine	Yes	> 20 mm	0.89 [0.52, 1.00]	1.00 [0.63, 1.00]
Gufler 1999	9	0	1	2	No	Supine	No	> 10 mm	0.90 [0.55, 1.00]	1.00 [0.16, 1.00]
Gufler 2004	4	0	0	3	No	Supine	No	> 10 mm	1.00 [0.40, 1.00]	1.00 [0.29, 1.00]
Healy 1997	4	1	0	5	No	Supine	Yes	> 0 mm	1.00 [0.40, 1.00]	0.83 [0.36, 1.00]
Kelvin 2000	9	0	0	1	Yes	Supine	Yes	> 0 mm	1.00 [0.66, 1.00]	1.00 [0.03, 1.00]
Lienemann 1997	10	2	1	27	Yes	Supine	Yes	> 30 mm	0.91 [0.59, 1.00]	0.93 [0.77, 0.99]
Martin 2017	34	0	2	2	Yes	Supine	Yes	> 0 mm	0.94 [0.81, 0.99]	1.00 [0.16, 1.00]
Matsuoka 2000	3	1	0	5	No	Prone	No	> 20 mm	1.00 [0.29, 1.00]	0.83 [0.36, 1.00]
Pilkington 2012	35	0	2	1	Yes	Supine	Yes	> 0 mm	0.95 [0.82, 0.99]	1.00 [0.03, 1.00]
Poncelet 2017	15	7	1	27	Yes	Supine	Yes	> 25 mm	0.94 [0.70, 1.00]	0.79 [0.62, 0.91]
Vanbeckevoort 1999	8	1	1	25	No	Supine	Yes	> 30 mm	0.89 [0.52, 1.00]	0.96 [0.80, 1.00]
Van Gruting 2017	58	25	2	37	Yes	Supine	Yes	> 20 mm	0.97 [0.88, 1.00]	0.60 [0.46, 0.72]
Van Iersel 2017	18	3	1	14	Yes	Supine	Yes	> 20 mm	0.95 [0.74, 1.00]	0.82 [0.57, 0.96]
Vitton 2011	42	1	4	9	Yes	Supine	Yes	> 20 mm	0.91 [0.79, 0.98]	0.90 [0.55, 1.00]
Zafar 2012	8	1	1	3	Yes	Supine	Yes	> 0 mm	0.89 [0.52, 1.00]	0.75 [0.19, 0.99]
Zafar 2017	12	9	1	31	Yes	Supine	Yes	> 20 mm	0.92 [0.64, 1.00]	0.78 [0.62, 0.89]

TPUS - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	16	1	1	15	Yes	Left-lateral	Yes	> 20 mm	0.94 [0.71, 1.00]	0.94 [0.70, 1.00]
Beer-Gabel 2015	21	11	3	70	No	Left-lateral	Yes	> 20 mm	0.88 [0.68, 0.97]	0.86 [0.77, 0.93]
Brusciano 2007	9	4	1	27	No	Supine	No	> 10 mm	0.90 [0.55, 1.00]	0.87 [0.70, 0.96]
Grasso 2007	30	1	4	8	No	Supine	No	> 0 mm	0.88 [0.73, 0.97]	0.89 [0.52, 1.00]
Hainsworth 2016	161	8	14	140	No	Supine	No	> 20 mm	0.92 [0.87, 0.96]	0.95 [0.90, 0.98]
Martellucci 2011	30	2	3	19	No	Supine	Part yes part no	> 10 mm	0.91 [0.76, 0.98]	0.90 [0.70, 0.99]
Perniola 2008	16	2	2	11	No	Supine	No	> 10 mm	0.89 [0.65, 0.99]	0.85 [0.55, 0.98]
Ron 2012	45	7	5	45	Yes	Left-lateral	Yes	> 20 mm	0.90 [0.78, 0.97]	0.87 [0.74, 0.94]
Steensma 2010	30	6	4	35	No	Supine	No	> 10 mm	0.88 [0.73, 0.97]	0.85 [0.71, 0.94]
Van Gruting 2017	54	13	11	53	No	Supine	No	> 10 mm	0.83 [0.72, 0.91]	0.80 [0.69, 0.89]
Weemhoff 2013	7	3	1	39	No	Supine	No	> 20 mm	0.88 [0.47, 1.00]	0.93 [0.81, 0.99]

EVUS - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	143	56	32	92	No	Supine	No	> 0 mm	0.82 [0.75, 0.87]	0.62 [0.54, 0.70]
Van Gruting 2017	30	6	35	60	No	Supine	No	> 10 mm	0.46 [0.34, 0.59]	0.91 [0.81, 0.97]

DAE - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	14	2	5	22	No	Left-lateral	Part yes part no	> 0 mm	0.74 [0.49, 0.91]	0.92 [0.73, 0.99]
Vitton 2011	36	1	10	9	No	Left-lateral	Yes	> 20 mm	0.78 [0.64, 0.89]	0.90 [0.55, 1.00]

EDF - Rectocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	18	1	1	4	Yes	Left-lateral	Yes	> 0 mm	0.95 [0.74, 1.00]	0.80 [0.28, 0.99]
Murad-Regadas 2008	23	1	1	5	No	Left-lateral	Yes	> 0 mm	0.96 [0.79, 1.00]	0.83 [0.36, 1.00]
Murad-Regadas 2011	22	0	1	6	No	Left-lateral	Yes	> 0 mm	0.96 [0.78, 1.00]	1.00 [0.54, 1.00]
Regadas 2011	75	1	2	8	No	Left-lateral	Yes	> 0 mm	0.97 [0.91, 1.00]	0.89 [0.52, 1.00]

Enterocele

Figure 6 shows the ROC plot and Figure 7 the forest plots for diagnosis of enterocele by all imaging techniques. The estimated pooled sensitivity of EP is 91% (CrI 83% to 97%), of MRI 85% (CrI 72% to 94%), TPUS 84% (CrI 63% to 96%), EVUS 68% (CrI 51% to 91%), DAE 74% (CrI 52% to 94%), and EDF 71% (CrI 51% to 96%) (Table 11). The probability that the sensitivity of the index test is equal to or better than EP is for all lower than 0.40 (Table 12). The estimated pooled specificity of EP is 96% (CrI 93% to 99%), of MRI 99% (CrI 96% to 100%), TPUS 98%, (CrI 95% to 100%), EVUS 97% (CrI 80% to 99%), DAE 97% (CrI 75% to 100%), and EDF 97% (CrI 87% to 100%). The probability that the specificity of index test is equal to or better than EP is higher than 0.40 for all. None meet the criteria for replacement

test nor for SnOUT triage test. All meet the criteria for SpIN triage test. In case of a positive test, the percentage of women that would truly have an enterocele is 89% for EP, 97% for MRI, 94% for TPUS, 87% for EVUS, 88% for DAE, and 90% for EDF, given a prevalence of 24%. In case of a negative test, 97% for EP, 95% for MRI, 95% for TPUS, 90% for EVUS, 92% for DAE, and 91% for EDF of women will truly not have an enterocele.

Figure 6. Summary ROC Plot from results of the LCA for enterocele for all imaging techniques. Note that the presented confidence regions are wider than the credibility regions reported from the LCA.

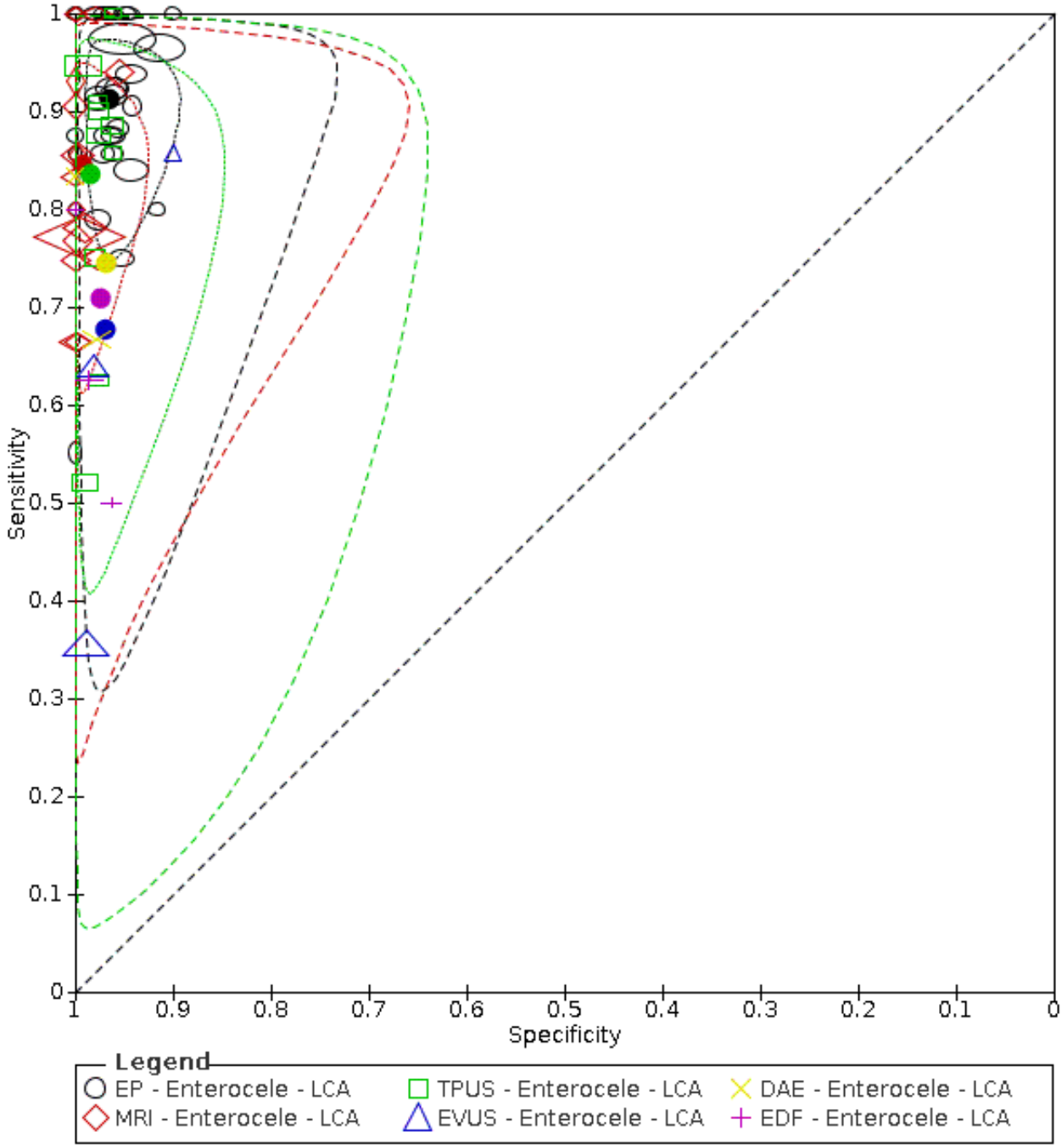
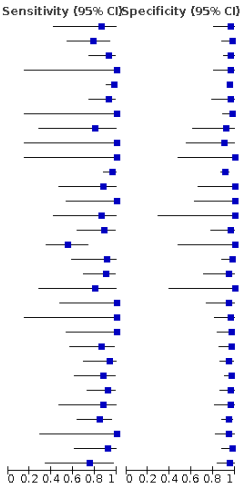


Figure 7. Forest plots of all tests for diagnosis of enterocele based on study specific results of the LCA. Note that these are different from the extracted data.

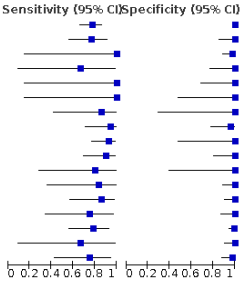
EP - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	6	1	1	25	Yes	Upright	Yes	Near RV septum	0.86 [0.42, 1.00]	0.96 [0.80, 1.00]
Beer-Gabel 2008	15	1	4	42	Yes	Upright	Yes	Near RV septum	0.79 [0.54, 0.94]	0.98 [0.88, 1.00]
Beer-Gabel 2015	24	3	2	76	Yes	Upright	Yes	Near RV septum	0.92 [0.75, 0.99]	0.96 [0.89, 0.99]
Brusciano 2007	2	1	0	25	Yes	Upright	Yes	Into RV space	1.00 [0.16, 1.00]	0.96 [0.80, 1.00]
Faggian 2013	74	26	2	512	Yes	Upright	Yes	Into RV space	0.97 [0.91, 1.00]	0.95 [0.93, 0.97]
Faucheron 2014	24	1	2	23	Yes	Upright	Yes	Into RV space	0.92 [0.75, 0.99]	0.96 [0.79, 1.00]
Fiaschetti 2013	2	1	0	46	Yes	Upright	Yes	> 10 mm below PCL	1.00 [0.16, 1.00]	0.98 [0.89, 1.00]
Foti 2013	4	1	1	11	Yes	Upright	Yes	> 0 mm below PCL	0.80 [0.28, 0.99]	0.92 [0.62, 1.00]
Gufler 1999	2	1	0	9	No	Upright	Yes	> 0 mm below PCL	1.00 [0.16, 1.00]	0.90 [0.55, 1.00]
Gufler 2004	2	0	0	5	No	Upright	Yes	> 0 mm below PCL	1.00 [0.16, 1.00]	1.00 [0.48, 1.00]
Hainsworth 2016	54	23	2	244	Yes	Upright	Yes	Near RV septum	0.96 [0.88, 1.00]	0.91 [0.87, 0.94]
Halligan 1996	7	0	1	9	Yes	Upright	Yes	Presence	0.88 [0.47, 1.00]	1.00 [0.66, 1.00]
Karaus 2000	6	0	0	8	Yes	Upright	Yes	Presence	1.00 [0.54, 1.00]	1.00 [0.63, 1.00]
Kelvin 2000	6	0	1	3	Yes	Upright	Yes	> 0 mm below PCL	0.86 [0.42, 1.00]	1.00 [0.29, 1.00]
Lienemann 1997	15	1	2	22	Yes	Upright	Yes	Into RV space	0.88 [0.64, 0.99]	0.96 [0.78, 1.00]
Lienemann 2000	16	0	13	5	Yes	Upright	Yes	> 0 mm below PCL	0.55 [0.36, 0.74]	1.00 [0.48, 1.00]
Martellucci 2011	10	1	1	42	Yes	Upright	Yes	Near RV septum	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]
Martin 2017	19	1	2	16	Yes	Upright	Yes	Into RV space	0.90 [0.70, 0.99]	0.94 [0.71, 1.00]
Matsuoka 2000	4	0	1	4	Yes	Upright	Yes	Presence	0.80 [0.28, 0.99]	1.00 [0.40, 1.00]
Miravalle 2016	5	1	0	18	Yes	Upright	Yes	> 0 mm below PCL	1.00 [0.48, 1.00]	0.95 [0.74, 1.00]
Murad-Regadas 2011	2	1	0	26	Yes	Upright	Yes	Below ischiococcygeal line	1.00 [0.16, 1.00]	0.96 [0.81, 1.00]
Pilkington 2012	6	1	0	31	Yes	Upright	Yes	Presence	1.00 [0.54, 1.00]	0.97 [0.84, 1.00]
Poncelet 2017	12	1	2	35	Yes	Upright	Yes	Extent into upper 1/2 of vagina	0.86 [0.57, 0.98]	0.97 [0.85, 1.00]
Regadas 2011	15	4	1	66	Yes	Upright	Yes	Below ischiococcygeal line	0.94 [0.70, 1.00]	0.94 [0.86, 0.98]
Ron 2012	14	3	2	83	Yes	Upright	Yes	Near RV septum	0.88 [0.62, 0.98]	0.97 [0.90, 0.99]
Steensma 2010	22	2	2	49	Yes	Upright	Yes	Extent into upper 1/3 of vagina	0.92 [0.73, 0.99]	0.96 [0.87, 1.00]
Vanbeckevoort 1999	7	1	1	26	Yes	Upright	Yes	> 0 mm below PCL	0.88 [0.47, 1.00]	0.96 [0.81, 1.00]
Van Gruting 2017	21	6	4	100	Yes	Upright	Yes	> 0 mm below PCL	0.84 [0.64, 0.95]	0.94 [0.88, 0.98]
Van Iersel 2017	3	2	0	36	Yes	Upright	Yes	> 0 mm below PCL	1.00 [0.29, 1.00]	0.95 [0.82, 0.99]
Vitton 2011	11	1	1	43	Yes	Upright	Yes	> 0 mm below PCL	0.92 [0.62, 1.00]	0.98 [0.88, 1.00]
Weemhoff 2013	6	2	2	40	Yes	Upright	Yes	Extent into upper 1/2 of vagina	0.75 [0.35, 0.97]	0.95 [0.84, 0.99]



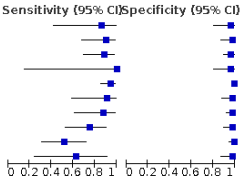
MRI - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Faggian 2013	58	1	17	538	Yes	Supine	Yes	Into RV space	0.77 [0.66, 0.86]	1.00 [0.99, 1.00]
Faucheron 2014	20	0	6	24	Yes	Supine	Yes	> 0 mm below PCL	0.77 [0.56, 0.91]	1.00 [0.86, 1.00]
Fiaschetti 2013	2	1	0	46	Yes	Supine	Yes	> 10 mm below PCL	1.00 [0.16, 1.00]	0.98 [0.89, 1.00]
Foti 2013	2	0	1	14	Yes	Supine	Yes	> 0 mm below PCL	0.67 [0.09, 0.99]	1.00 [0.77, 1.00]
Gufler 1999	2	0	0	10	No	Supine	No	> 0 mm below PCL	1.00 [0.16, 1.00]	1.00 [0.69, 1.00]
Gufler 2004	2	0	0	5	No	Supine	No	> 0 mm below PCL	1.00 [0.16, 1.00]	1.00 [0.48, 1.00]
Kelvin 2000	6	0	1	3	Yes	Supine	Yes	> 0 mm below PCL	0.86 [0.42, 1.00]	1.00 [0.29, 1.00]
Lienemann 1997	16	1	1	22	Yes	Supine	Yes	> 0 mm below PCL	0.94 [0.71, 1.00]	0.96 [0.78, 1.00]
Lienemann 2000	27	0	2	5	Yes	Supine	Yes	> 0 mm below PCL	0.93 [0.77, 0.99]	1.00 [0.48, 1.00]
Martin 2017	19	0	2	17	Yes	Supine	Yes	Into RV space	0.90 [0.70, 0.99]	1.00 [0.80, 1.00]
Matsuoka 2000	4	0	1	4	No	Prone	No	Presence	0.80 [0.28, 0.99]	1.00 [0.40, 1.00]
Pilkington 2012	5	0	1	32	Yes	Supine	Yes	Presence	0.83 [0.36, 1.00]	1.00 [0.89, 1.00]
Poncelet 2017	12	0	2	36	Yes	Supine	Yes	Extent into upper 1/2 of vagina	0.86 [0.57, 0.98]	1.00 [0.90, 1.00]
Vanbeckevoort 1999	6	0	2	27	No	Supine	Yes	> 0 mm below PCL	0.75 [0.35, 0.97]	1.00 [0.87, 1.00]
Van Gruting 2017	18	1	5	98	Yes	Supine	Yes	> 0 mm below PCL	0.78 [0.56, 0.93]	0.99 [0.95, 1.00]
Van Iersel 2017	2	0	1	38	Yes	Supine	Yes	> 0 mm below PCL	0.67 [0.09, 0.99]	1.00 [0.91, 1.00]
Vitton 2011	9	1	3	43	Yes	Supine	Yes	> 0 mm below PCL	0.75 [0.43, 0.95]	0.98 [0.88, 1.00]



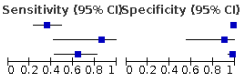
TPUS - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	6	1	1	25	Yes	Left-lateral	Yes	Near RV septum	0.86 [0.42, 1.00]	0.96 [0.80, 1.00]
Beer-Gabel 2008	18	1	2	41	In some	Left-lateral	Yes	Near RV septum	0.90 [0.68, 0.99]	0.98 [0.87, 1.00]
Beer-Gabel 2015	23	3	3	76	No	Left-lateral	Yes	Near RV septum	0.88 [0.70, 0.98]	0.96 [0.89, 0.99]
Brusciano 2007	2	1	0	25	No	Supine	No	Into RV space	1.00 [0.16, 1.00]	0.96 [0.80, 1.00]
Hainsworth 2016	53	2	3	265	No	Supine	No	Extent into upper 1/3 of vagina	0.95 [0.85, 0.99]	0.99 [0.97, 1.00]
Martellucci 2011	10	1	1	42	No	Supine	Part yes part no	Extent into upper 1/3 of vagina	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]
Ron 2012	14	2	2	84	Yes	Left-lateral	Yes	Near RV septum	0.88 [0.62, 0.98]	0.98 [0.92, 1.00]
Steensma 2010	18	1	6	50	No	Supine	No	Extent into upper 1/3 of vagina	0.75 [0.53, 0.90]	0.98 [0.90, 1.00]
Van Gruting 2017	13	1	12	105	No	Supine	No	Below PB	0.52 [0.31, 0.72]	0.99 [0.95, 1.00]
Weemhoff 2013	5	1	3	41	No	Supine	No	Extent into upper 1/2 of vagina	0.63 [0.24, 0.91]	0.98 [0.87, 1.00]



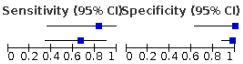
EVUS - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	20	3	36	264	No	Supine	No	Near RV septum	0.36 [0.23, 0.50]	0.99 [0.97, 1.00]
Halligan 1996	6	1	1	9	No	Left-lateral	No	Near RV septum	0.86 [0.42, 1.00]	0.90 [0.55, 1.00]
Van Gruting 2017	16	2	9	104	No	Supine	No	Near RV septum	0.64 [0.43, 0.82]	0.98 [0.93, 1.00]



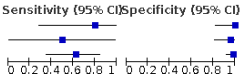
DAE - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Karaus 2000	5	0	1	8	No	Left-lateral	No	Presence	0.83 [0.36, 1.00]	1.00 [0.63, 1.00]
Vitton 2011	8	1	4	43	No	Left-lateral	Yes	Into RV space	0.67 [0.35, 0.90]	0.98 [0.88, 1.00]



EDF - Enterocele - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	4	0	1	19	Yes	Left-lateral	Yes	Below ischiococcygeal line	0.80 [0.28, 0.99]	1.00 [0.82, 1.00]
Murad-Regadas 2011	1	1	1	26	No	Left-lateral	Yes	Presence	0.50 [0.01, 0.99]	0.96 [0.81, 1.00]
Regadas 2011	10	1	6	69	No	Left-lateral	Yes	> 0 mm below PCL	0.63 [0.35, 0.85]	0.99 [0.92, 1.00]



Intussusception

Figure 8 shows the ROC plot and Figure 9 the forest plots for diagnosis of intussusception on all imaging techniques. The estimated pooled sensitivity of EP is 89% (CrI 79% to 96%), of MRI 61% (CrI 51% to 78%), TPUS 75% (CrI 54% to 93%), EVUS 63% (CrI 51% to 88%), DAE 61% (CrI 50% to 89%), and EDF 89% (CrI 65% to 98%) (Table 11). The probability that the sensitivity of the index test is equal to or better than EP is for all lower than 0.40, except for EDF 0.52 (Table 12). The estimated pooled specificity of EP is 92% (CrI 86% to 97%), of MRI 97% (CrI 88% to 100%), TPUS 96% (CrI 91% to 99%), EVUS 93% (CrI 72% to 99%), DAE 93% (CrI 65% to 99%), and EDF 92% (CrI 72% to 99%). The probability that the specificity of the index test is equal to or better than EP is higher than 0.40 for all. Only EDF meets the criteria of a replacement test and SnOUT triage test. All meet the criteria for SpIN triage test. In case of a positive test, the percentage of women that would truly have an intussusception is 89% for EP, 94% for MRI, 94% for TPUS, 87% for EVUS, 87% for DAE, and 90% for EDF, given a prevalence of 44%. In case of a negative test, 91% for EP, 76% for MRI, 83% for TPUS, 76% for EVUS, 75% for DAE, and 92% for EDF of women will truly not have an intussusception.

Figure 8. Summary ROC Plot from results of the LCA for intussusception for all imaging techniques. Note that the presented confidence regions are wider than the credibility regions reported from the LCA.

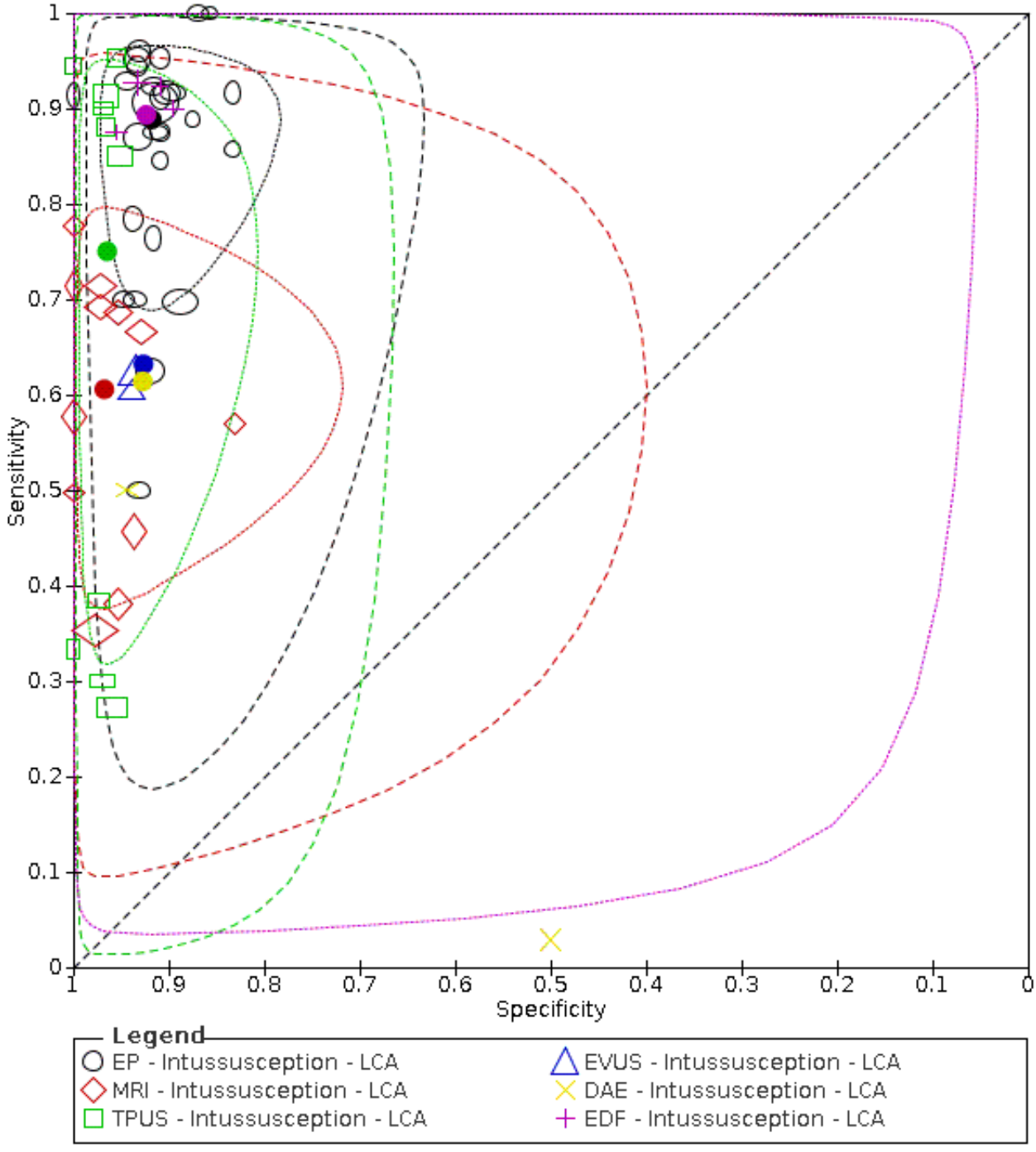
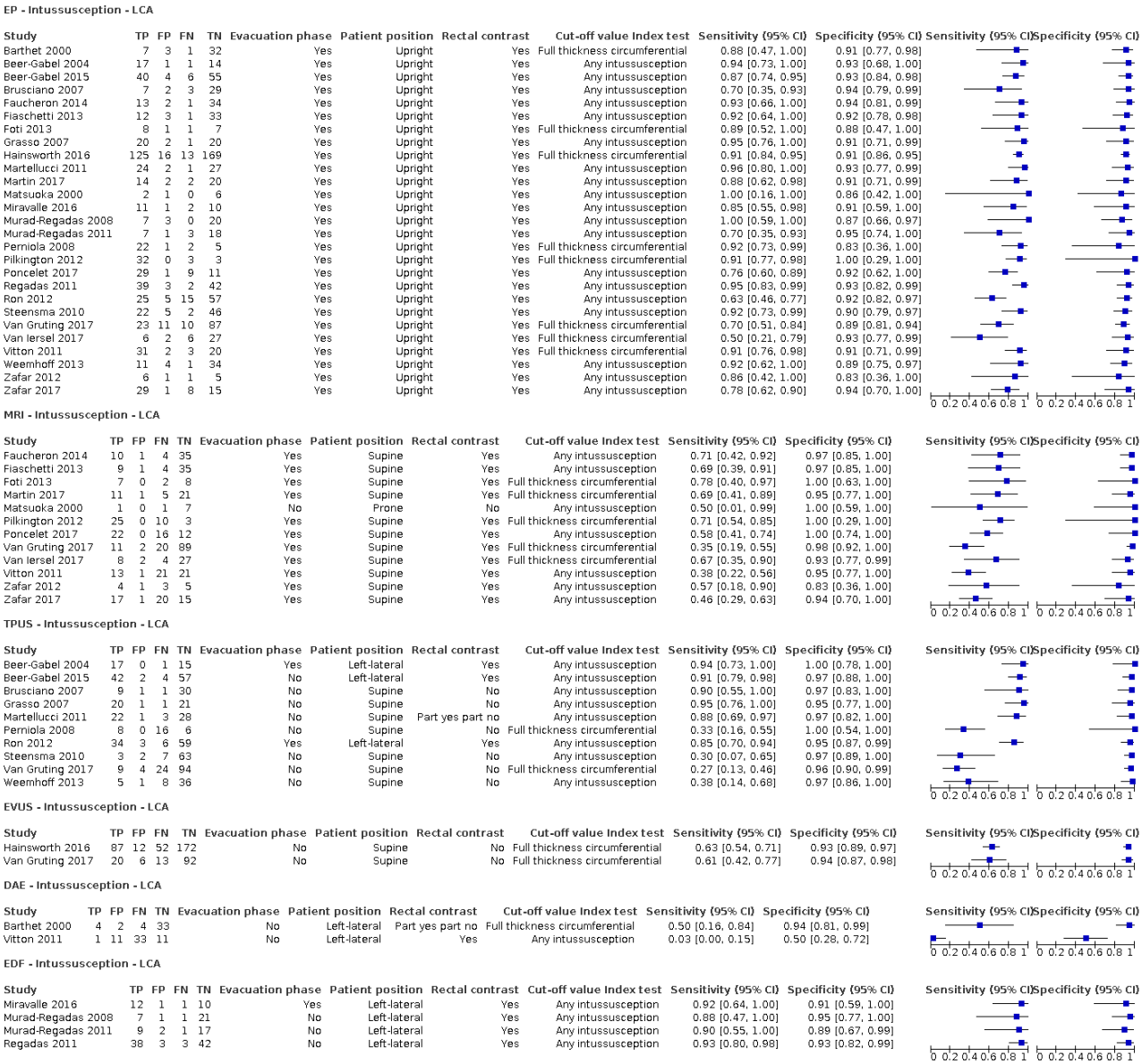


Figure 9. Forest plots of all tests for diagnosis of intussusception based on study specific results of the LCA. Note that these are different from the extracted data.



EP - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	7	3	1	32	Yes	Upright	Yes	Full thickness circumferential	0.88 [0.47, 1.00]	0.91 [0.77, 0.98]
Beer-Gabel 2004	17	1	1	14	Yes	Upright	Yes	Any intussusception	0.94 [0.73, 1.00]	0.93 [0.68, 1.00]
Beer-Gabel 2015	40	4	6	55	Yes	Upright	Yes	Any intussusception	0.87 [0.74, 0.95]	0.93 [0.84, 0.98]
Brusciano 2007	7	2	3	29	Yes	Upright	Yes	Any intussusception	0.70 [0.35, 0.93]	0.94 [0.79, 0.99]
Faucheron 2014	13	2	1	34	Yes	Upright	Yes	Any intussusception	0.93 [0.66, 1.00]	0.94 [0.81, 0.99]
Fiaschetti 2013	12	3	1	33	Yes	Upright	Yes	Any intussusception	0.92 [0.64, 1.00]	0.92 [0.78, 0.98]
Foti 2013	8	1	1	7	Yes	Upright	Yes	Full thickness circumferential	0.89 [0.52, 1.00]	0.88 [0.47, 1.00]
Grasso 2007	20	2	1	20	Yes	Upright	Yes	Any intussusception	0.95 [0.76, 1.00]	0.91 [0.71, 0.99]
Hainsworth 2016	125	16	13	169	Yes	Upright	Yes	Full thickness circumferential	0.91 [0.84, 0.95]	0.91 [0.86, 0.95]
Martellucci 2011	24	2	1	27	Yes	Upright	Yes	Any intussusception	0.96 [0.80, 1.00]	0.93 [0.77, 0.99]
Martin 2017	14	2	2	20	Yes	Upright	Yes	Any intussusception	0.88 [0.62, 0.98]	0.91 [0.71, 0.99]
Matsuoka 2000	2	1	0	6	Yes	Upright	Yes	Any intussusception	1.00 [0.16, 1.00]	0.86 [0.42, 1.00]
Miravalle 2016	11	1	2	10	Yes	Upright	Yes	Any intussusception	0.85 [0.55, 0.98]	0.91 [0.59, 1.00]
Murad-Regadas 2008	7	3	0	20	Yes	Upright	Yes	Any intussusception	1.00 [0.59, 1.00]	0.87 [0.66, 0.97]
Murad-Regadas 2011	7	1	3	18	Yes	Upright	Yes	Any intussusception	0.70 [0.35, 0.93]	0.95 [0.74, 1.00]
Perniola 2008	22	1	2	5	Yes	Upright	Yes	Full thickness circumferential	0.92 [0.73, 0.99]	0.83 [0.36, 1.00]
Pilkington 2012	32	0	3	3	Yes	Upright	Yes	Full thickness circumferential	0.91 [0.77, 0.98]	1.00 [0.29, 1.00]
Poncelet 2017	29	1	9	11	Yes	Upright	Yes	Any intussusception	0.76 [0.60, 0.89]	0.92 [0.62, 1.00]
Regadas 2011	39	3	2	42	Yes	Upright	Yes	Any intussusception	0.95 [0.83, 0.99]	0.93 [0.82, 0.99]
Ron 2012	25	5	15	57	Yes	Upright	Yes	Any intussusception	0.63 [0.46, 0.77]	0.92 [0.82, 0.97]
Steensma 2010	22	5	2	46	Yes	Upright	Yes	Any intussusception	0.92 [0.73, 0.99]	0.90 [0.79, 0.97]
Van Gruting 2017	23	11	10	87	Yes	Upright	Yes	Full thickness circumferential	0.70 [0.51, 0.84]	0.89 [0.81, 0.94]
Van Iersel 2017	6	2	6	27	Yes	Upright	Yes	Full thickness circumferential	0.50 [0.21, 0.79]	0.93 [0.77, 0.99]
Vitton 2011	31	2	3	20	Yes	Upright	Yes	Full thickness circumferential	0.91 [0.76, 0.98]	0.91 [0.71, 0.99]
Weemhoff 2013	11	4	1	34	Yes	Upright	Yes	Any intussusception	0.92 [0.62, 1.00]	0.89 [0.75, 0.97]
Zafar 2012	6	1	1	5	Yes	Upright	Yes	Any intussusception	0.86 [0.42, 1.00]	0.83 [0.36, 1.00]
Zafar 2017	29	1	8	15	Yes	Upright	Yes	Any intussusception	0.78 [0.62, 0.90]	0.94 [0.70, 1.00]

MRI - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Faucheron 2014	10	1	4	35	Yes	Supine	Yes	Any intussusception	0.71 [0.42, 0.92]	0.97 [0.85, 1.00]
Fiaschetti 2013	9	1	4	35	Yes	Supine	Yes	Any intussusception	0.69 [0.39, 0.91]	0.97 [0.85, 1.00]
Foti 2013	7	0	2	8	Yes	Supine	Yes	Full thickness circumferential	0.78 [0.40, 0.97]	1.00 [0.63, 1.00]
Martin 2017	11	1	5	21	Yes	Supine	Yes	Full thickness circumferential	0.69 [0.41, 0.89]	0.95 [0.77, 1.00]
Matsuoka 2000	1	0	1	7	No	Prone	No	Any intussusception	0.50 [0.01, 0.99]	1.00 [0.59, 1.00]
Pilkington 2012	25	0	10	3	Yes	Supine	Yes	Full thickness circumferential	0.71 [0.54, 0.85]	1.00 [0.29, 1.00]
Poncelet 2017	22	0	16	12	Yes	Supine	Yes	Any intussusception	0.58 [0.41, 0.74]	1.00 [0.74, 1.00]
Van Gruting 2017	11	2	20	89	Yes	Supine	Yes	Full thickness circumferential	0.35 [0.19, 0.55]	0.98 [0.92, 1.00]
Van Iersel 2017	8	2	4	27	Yes	Supine	Yes	Full thickness circumferential	0.67 [0.35, 0.90]	0.93 [0.77, 0.99]
Vitton 2011	13	1	21	21	Yes	Supine	Yes	Any intussusception	0.38 [0.22, 0.56]	0.95 [0.77, 1.00]
Zafar 2012	4	1	3	5	Yes	Supine	Yes	Any intussusception	0.57 [0.18, 0.90]	0.83 [0.36, 1.00]
Zafar 2017	17	1	20	15	Yes	Supine	Yes	Any intussusception	0.46 [0.29, 0.63]	0.94 [0.70, 1.00]

TPUS - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	17	0	1	15	Yes	Left-lateral	Yes	Any intussusception	0.94 [0.73, 1.00]	1.00 [0.78, 1.00]
Beer-Gabel 2015	42	2	4	57	No	Left-lateral	Yes	Any intussusception	0.91 [0.79, 0.98]	0.97 [0.88, 1.00]
Brusciano 2007	9	1	1	30	No	Supine	No	Any intussusception	0.90 [0.55, 1.00]	0.97 [0.83, 1.00]
Grasso 2007	20	1	1	21	No	Supine	No	Any intussusception	0.95 [0.76, 1.00]	0.95 [0.77, 1.00]
Martellucci 2011	22	1	3	28	No	Supine	Part yes part no	Any intussusception	0.88 [0.69, 0.97]	0.97 [0.82, 1.00]
Perniola 2008	8	0	16	6	No	Supine	No	Full thickness circumferential	0.33 [0.16, 0.55]	1.00 [0.54, 1.00]
Ron 2012	34	3	6	59	Yes	Left-lateral	Yes	Any intussusception	0.85 [0.70, 0.94]	0.95 [0.87, 0.99]
Steensma 2010	3	2	7	63	No	Supine	No	Any intussusception	0.30 [0.07, 0.65]	0.97 [0.89, 1.00]
Van Gruting 2017	9	4	24	94	No	Supine	No	Full thickness circumferential	0.27 [0.13, 0.46]	0.96 [0.90, 0.99]
Weemhoff 2013	5	1	8	36	No	Supine	No	Any intussusception	0.38 [0.14, 0.68]	0.97 [0.86, 1.00]

EVUS - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	87	12	52	172	No	Supine	No	Full thickness circumferential	0.63 [0.54, 0.71]	0.93 [0.89, 0.97]
Van Gruting 2017	20	6	13	92	No	Supine	No	Full thickness circumferential	0.61 [0.42, 0.77]	0.94 [0.87, 0.98]

DAE - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	4	2	4	33	No	Left-lateral	Part yes part no	Full thickness circumferential	0.50 [0.16, 0.84]	0.94 [0.81, 0.99]
Vitton 2011	1	11	33	11	No	Left-lateral	Yes	Any intussusception	0.03 [0.00, 0.15]	0.50 [0.28, 0.72]

EDF - Intussusception - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	12	1	1	10	Yes	Left-lateral	Yes	Any intussusception	0.92 [0.64, 1.00]	0.91 [0.59, 1.00]
Murad-Regadas 2008	7	1	1	21	No	Left-lateral	Yes	Any intussusception	0.88 [0.47, 1.00]	0.95 [0.77, 1.00]
Murad-Regadas 2011	9	2	1	17	No	Left-lateral	Yes	Any intussusception	0.90 [0.55, 1.00]	0.89 [0.67, 0.99]
Regadas 2011	38	3	3	42	No	Left-lateral	Yes	Any intussusception	0.93 [0.80, 0.98]	0.93 [0.82, 0.99]

Anismus

Figure 10 shows the ROC plot and Figure 11 the forest plots for diagnosis of anismus on all imaging techniques. The estimated pooled sensitivity of EP is 80% (CrI 63% to 94%), of MRI 86% (CrI 60% to 98%), TPUS 92% (CrI 72% to 98%), EVUS 84% (CrI 59% to 96%), and EDF 87% (CrI 72% to 96%) (Table 11). The probability that the sensitivity of the index test is equal to or better than EP is for all higher than 0.40 (Table 12). The estimated pooled specificity of EP is 97% (CrI 94% to 99%), of MRI 96% (CrI 89% to 99%), TPUS 91% (CrI 83% to 97%), EVUS 90% (CrI 63% to 98%), and EDF 93% (CrI 74% to 99%). The probability that the specificity of the index test is equal to or better than EP is lower than 0.40 for all. None meet the criteria for a replacement test nor for SpIN triage test. All meet the criteria for SnOUT triage test. In case of a positive test, the percentage of women that would truly have anismus is 89% for EP, 87% for MRI, 77% for TPUS, 74% for EVUS, and 80% for EDF, given a prevalence of 25%. In case of a negative test, 94% for EP, 95% for MRI, 97% for TPUS, 95% for EVUS, and 96% for EDF of women will truly not have anismus. No results for DAE are available as no studies have performed DAE for diagnosis of anismus.

Figure 10. Summary ROC Plot from results of the LCA for anismus for all imaging techniques. Note that the presented confidence regions are wider than the credibility regions reported from the LCA.

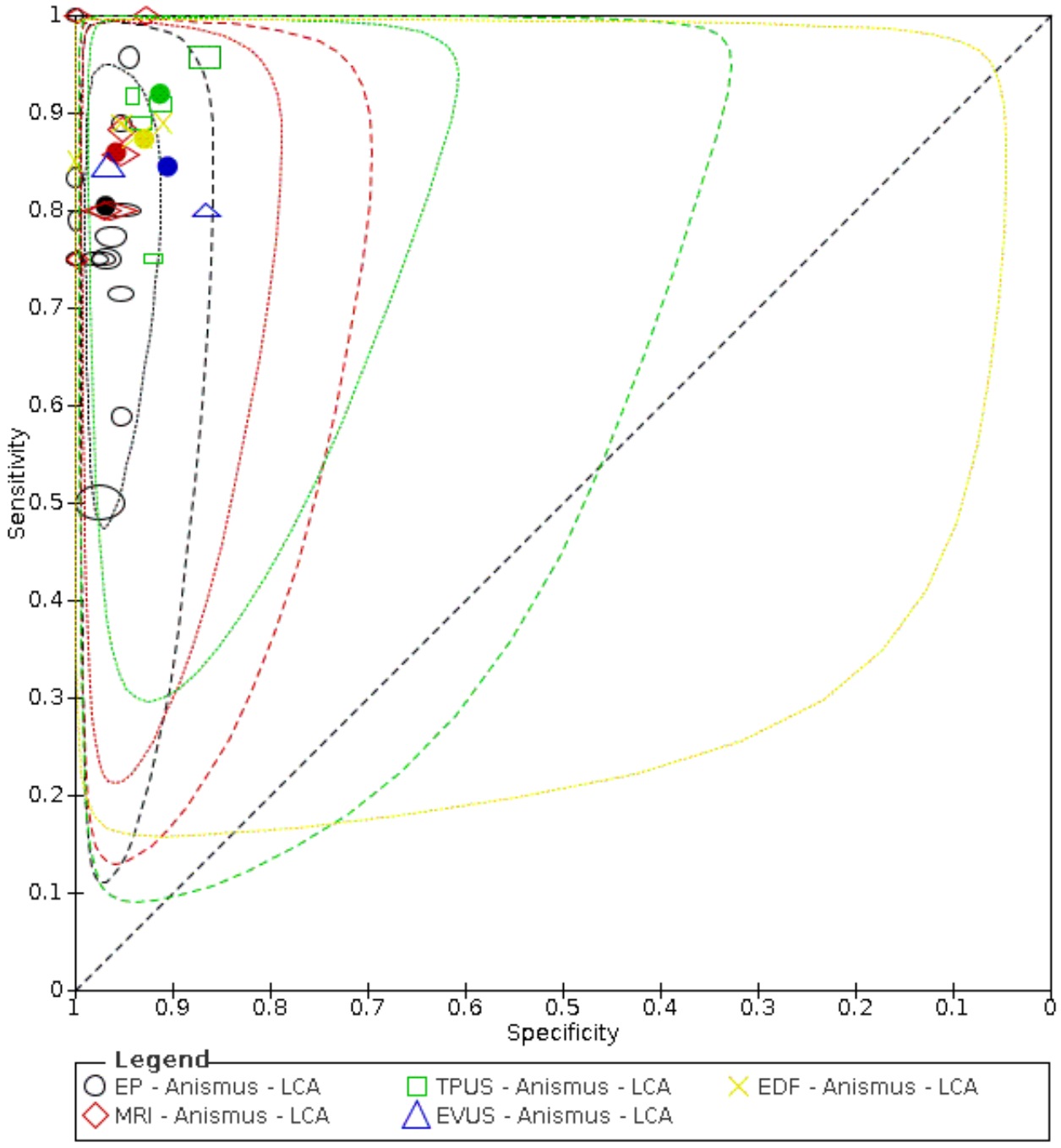
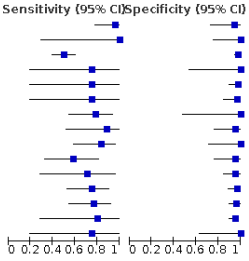


Figure 11. Forest plots of all tests for diagnosis of anismus based on study specific results of the LCA. Note that these are different from the extracted data.

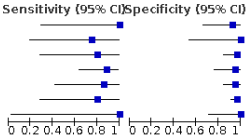
EP - Anismus - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Brusciano 2007	22	1	1	17	Yes	Upright	Yes	Presence	0.96 [0.78, 1.00]	0.94 [0.73, 1.00]
Foti 2013	3	0	0	13	Yes	Upright	Yes	ARA more acute	1.00 [0.29, 1.00]	1.00 [0.75, 1.00]
Hainsworth 2016	46	6	46	226	Yes	Upright	Yes	Paradoxical contraction	0.50 [0.39, 0.61]	0.97 [0.94, 0.99]
Healy 1997	3	0	1	6	Yes	Upright	Yes	ARA more acute	0.75 [0.19, 0.99]	1.00 [0.54, 1.00]
Martellucci 2011	3	1	1	49	Yes	Upright	Yes	ARA more acute	0.75 [0.19, 0.99]	0.98 [0.89, 1.00]
Martin 2017	3	1	1	33	Yes	Upright	Yes	ARA more acute	0.75 [0.19, 0.99]	0.97 [0.85, 1.00]
Miravalle 2016	15	0	4	5	Yes	Upright	Yes	ARA more acute	0.79 [0.54, 0.94]	1.00 [0.48, 1.00]
Murad-Regadas 2008	8	1	1	20	Yes	Upright	Yes	ARA more acute	0.89 [0.52, 1.00]	0.95 [0.76, 1.00]
Murad-Regadas 2011	15	0	3	11	Yes	Upright	Yes	ARA more acute	0.83 [0.59, 0.96]	1.00 [0.72, 1.00]
Pilkington 2012	10	1	7	20	Yes	Upright	Yes	Paradoxical contraction	0.59 [0.33, 0.82]	0.95 [0.76, 1.00]
Poncelet 2017	5	2	2	41	Yes	Upright	Yes	Paradoxical contraction	0.71 [0.29, 0.96]	0.95 [0.84, 0.99]
Regadas 2011	18	2	6	60	Yes	Upright	Yes	ARA more acute	0.75 [0.53, 0.90]	0.97 [0.89, 1.00]
Ron 2012	17	3	5	77	Yes	Upright	Yes	Presence	0.77 [0.55, 0.92]	0.96 [0.89, 0.99]
Van Gruting 2017	4	6	1	120	Yes	Upright	Yes	Paradoxical contraction	0.80 [0.28, 0.99]	0.95 [0.90, 0.98]
Zafar 2012	3	0	1	8	Yes	Upright	Yes	Paradoxical contraction	0.75 [0.19, 0.99]	1.00 [0.63, 1.00]



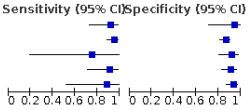
MRI - Anismus - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Foti 2013	3	1	0	13	Yes	Supine	Yes	ARA more acute	1.00 [0.29, 1.00]	0.93 [0.66, 1.00]
Healy 1997	3	0	1	6	No	Supine	Yes	ARA more acute	0.75 [0.19, 0.99]	1.00 [0.54, 1.00]
Martin 2017	4	1	1	32	Yes	Supine	Yes	Paradoxical contraction	0.80 [0.28, 0.99]	0.97 [0.84, 1.00]
Pilkington 2012	15	1	2	20	Yes	Supine	Yes	Paradoxical contraction	0.88 [0.64, 0.99]	0.95 [0.76, 1.00]
Poncelet 2017	6	2	1	41	Yes	Supine	Yes	Paradoxical contraction	0.86 [0.42, 1.00]	0.95 [0.84, 0.99]
Van Gruting 2017	4	4	1	113	Yes	Supine	Yes	Paradoxical contraction	0.80 [0.28, 0.99]	0.97 [0.91, 0.99]
Zafar 2012	1	0	0	11	Yes	Supine	Yes	Paradoxical contraction	1.00 [0.03, 1.00]	1.00 [0.72, 1.00]



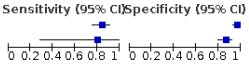
TPUS - Anismus - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Brusciano 2007	22	1	2	16	No	Supine	No	Presence	0.92 [0.73, 0.99]	0.94 [0.71, 1.00]
Hainsworth 2016	87	31	4	201	No	Supine	No	Paradoxical contraction	0.96 [0.89, 0.99]	0.87 [0.82, 0.91]
Martellucci 2011	3	4	1	46	No	Supine	Part yes part no	ARA more acute	0.75 [0.19, 0.99]	0.92 [0.81, 0.98]
Ron 2012	20	7	2	73	Yes	Left-lateral	Yes	Presence	0.91 [0.71, 0.99]	0.91 [0.83, 0.96]
Van Gruting 2017	8	8	1	114	No	Supine	No	Paradoxical contraction	0.89 [0.52, 1.00]	0.93 [0.87, 0.97]



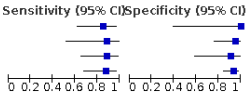
EVUS - Anismus - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	77	8	14	224	No	Supine	No	Paradoxical contraction	0.85 [0.76, 0.91]	0.97 [0.93, 0.98]
Van Gruting 2017	4	17	1	109	No	Supine	No	Paradoxical contraction	0.80 [0.28, 0.99]	0.87 [0.79, 0.92]



EDF - Anismus - LCA

Study	TP	FP	FN	TN	Evacuation phase	Patient position	Rectal contrast	Cut-off value Index test	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	17	0	3	4	Yes	Left-lateral	Yes	ARA more acute	0.85 [0.62, 0.97]	1.00 [0.40, 1.00]
Murad-Regadas 2008	8	1	1	20	No	Left-lateral	Yes	ARA more acute	0.89 [0.52, 1.00]	0.95 [0.76, 1.00]
Murad-Regadas 2011	16	1	2	10	No	Left-lateral	Yes	ARA more acute	0.89 [0.65, 0.99]	0.91 [0.59, 1.00]
Regadas 2011	21	4	3	58	No	Left-lateral	Yes	ARA more acute	0.88 [0.68, 0.97]	0.94 [0.84, 0.98]



Pelvic floor descent

Figure 12 shows the ROC plot and Figure 13 the forest plots for diagnosis of pelvic floor descent (PFD) on all imaging techniques. The estimated sensitivity of EP is 98% (CrI 93% to 100%), of MRI 94% (CrI 81% to 98%) and DAE 93% (CrI 64% to 99%) (Table 11). The probability that the sensitivity of the index test is equal to or better than EP is for all lower than 0.40 (Table 12). The estimated specificity of EP is 83% (CrI 59% to 96%), of MRI 79% (CrI 54% to 97%) and DAE 74% (CrI 54% to 93%). The probability that the specificity of the index test is equal to or better than EP is higher than 0.40 for MRI, but not for DAE. None meet the criteria for a replacement test nor SnOUT triage test. MRI meets the criteria for SpIN triage test. In case of a positive test, the percentage of women that would truly have PFD is 92% for EP, 90% for MRI, and 88% for DAE, given a prevalence of 67%. In case of a negative test, 94% for EP, 86% for MRI, and 84% for DAE of women will truly not have PFD. No results for TPUS, EVUS and EDF are available, as none or just one study was performed for diagnoses of PFD.

Figure 12. Summary ROC Plot from results of the LCA for pelvic floor descent for all imaging techniques. Note that the presented confidence regions are wider than the credibility regions reported from the LCA.

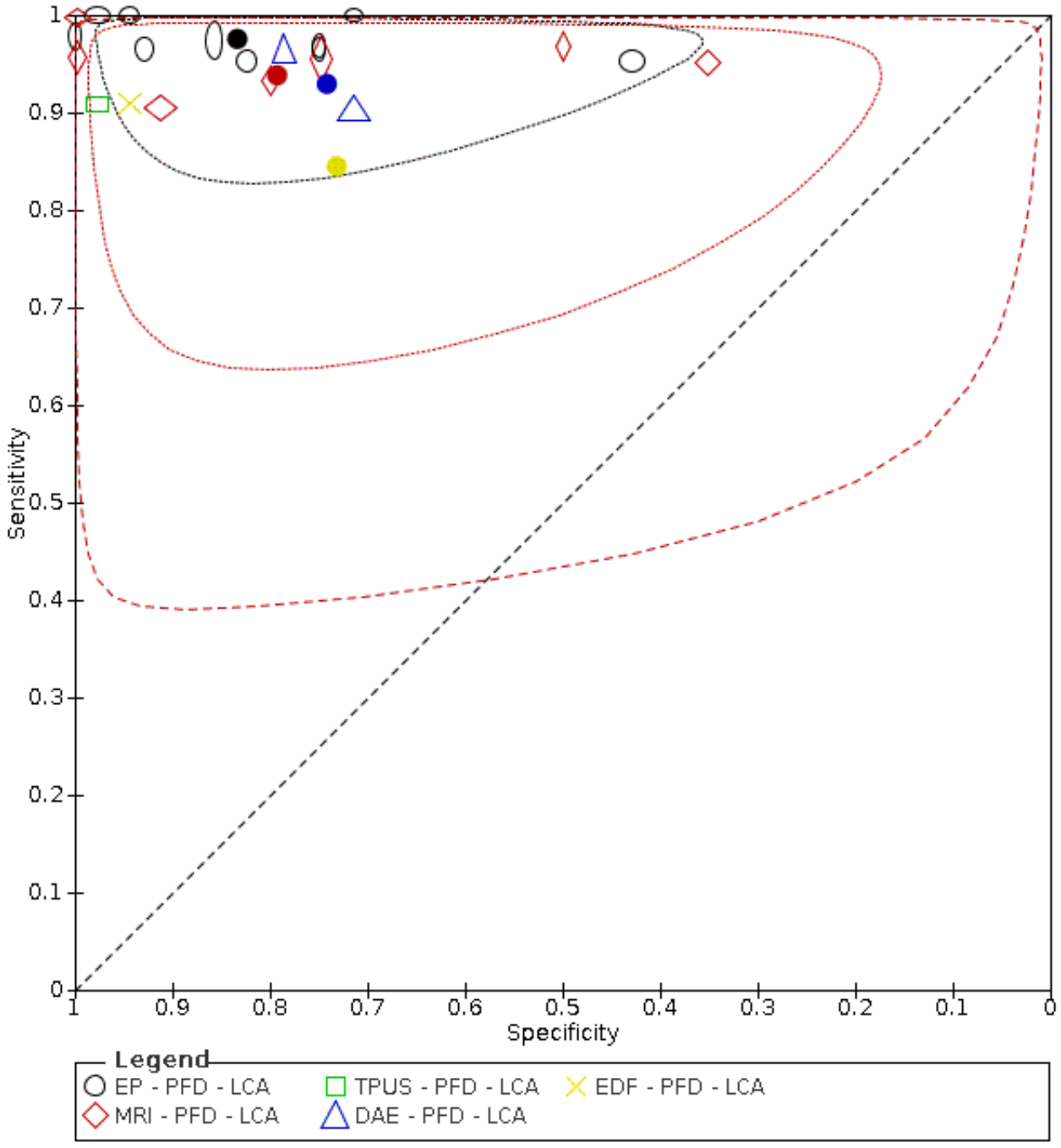
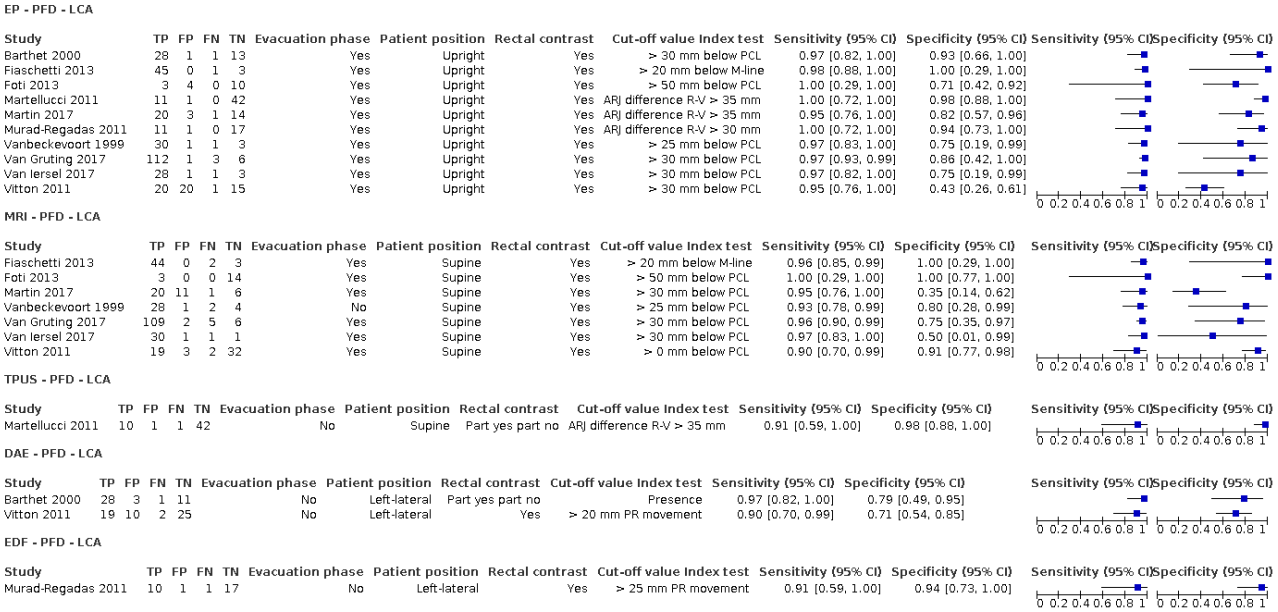


Figure 13. Forest plots of all tests for diagnosis of pelvic floor descent based on study specific results of the LCA. Note that these are different from the extracted data.



Heterogeneity analysis

For the analysis of possible sources of heterogeneity, some studies provided extra test accuracy data on additional test conditions; there was one study for the effect of evacuation phase (Foti 2013), five for the cut-off value of rectocele on MRI (Fiaschetti 2013; Healy 1997; Kelvin 2000; Martin 2017; Van Gruting 2017), and three for the cut-off value of rectocele on TPUS (Grasso 2007; Martellucci 2011; Steensma 2010). Heterogeneity analyses are performed for EP, MRI and TPUS, but not for EVUS, DAE and EFD because of low numbers of studies.

A. Effect of rectal contrast on test accuracy

The effect of rectal contrast on test accuracy is only assessed for TPUS. EP is performed with rectal contrast in all studies, so no heterogeneity analysis is necessary. For MRI this assessment could not be performed independently from evacuation phase, as in most studies that used rectal contrast. Also an evacuation phase was applied, and in all studies without rectal contrast no evacuation phase was possible.

Results for TPUS are presented in Table 13. Sensitivities of TPUS are all higher with rectal contrast for rectocele than without (92% versus 81%), enterocele (90% versus 67%) and intussusception (90% versus 61%). Specificities of TPUS are similar or lower with rectal contrast for rectocele than without (87% versus 88%), enterocele (95% versus 99%) and intussusception (90% versus 96%). The probability that the sensitivity of TPUS with rectal contrast is equal to or better than EP remained less than 0.40. The probability that specificity of TPUS with or without evacuation phase is equal to or better than EP remained more than 0.40. Studies that used rectal contrast also performed ultrasound in the left-lateral position and often with evacuation phase, which could also affect test accuracy results, but this could not be assessed independently.

B. Effect of evacuation phase on test accuracy

The effect of an evacuation phase on test accuracy is only assessed for MRI. In nearly all studies EP is performed with evacuation phase, except for Gufler 1999 and Gufler 2004, so no heterogeneity analysis was necessary. The effect of an evacuation phase on test accuracy of TPUS is not assessed independently, as all these studies also used rectal contrast.

Results for MRI are presented in Table 14. The sensitivities of MRI are higher with than without evacuation phase for rectocele (94% versus 65%) and enterocele (87% versus 62%). The specificities of MRI are similar to or lower with than without evacuation phase for rectocele (84% versus 95%) and enterocele (99% versus 97%). The probability that the sensitivity of MRI with evacuation phase is equal to or better than EP remained less than 0.40, and is very low for MRI without evacuation phase (probability 0.001 and 0.013). The probability that specificity of MRI with or without evacuation phase is equal to or better than EP remained above 0.40. For intussusception and anismus no differences in sensitivity and specificity of MRI with or without evacuation phase are observed. For the analysis of PFD the analyses did not converge.

C. Effect of participant position on test accuracy

The effect of participant position on test accuracy is not analysed. In all studies EP is performed with the participant in the upright position, so no heterogeneity analysis was necessary. For MRI this assessment was not possible as most studies performed MRI in the supine position. Only one study performed MRI in the upright position (Fiaschetti 2013), and two performed MRI in the prone position (Dellemare 1994; Matsuoka 2000). The effect of participant position on test accuracy of TPUS is not assessed independently, as all these studies also used rectal contrast.

D. Effect of cut-off value on test accuracy

The results of the analysis of the effect of cut-off values on test accuracy are presented in Table 15. When using ≥ 2 cm depth as the cut-off value for the diagnosis of rectocele, sensitivities of EP and MRI are lower (97% versus 99% and 93% versus 98% respectively) and specificities are higher (89% versus 55% and 94% versus 66% respectively) compared to the > 0 cm cut-off. For TPUS both sensitivity (91% versus 83%) and specificity (91% versus 69%) increase by taking the higher cut-off. For diagnosis of enterocele, the cut-off 'small bowel below the PCL' gives for both EP and MRI similar sensitivity and specificity compared to 'small bowel into the rectovaginal space'. For diagnosis of intussusception, full-thickness circumferential intussusception compared to any intussusception has a higher sensitivity for MRI (70% versus 58%), but a lower sensitivity for EP (83% versus 93%) and TPUS (61% versus 88%). There is no notable difference between the specificities using the two cut-off values for intussusception for all three imaging techniques. For diagnosis of anismus, measuring ARA has a higher sensitivity compared to paradoxical pelvic contraction for EP (85% versus 55%) and MRI (91% versus 70%). Specificities for both cut-off values for anismus are similar. Cut-off analysis for PFD could not be performed because of insufficient data.

Sensitivity analyses

1. Excluding studies at high risk of bias

Compared to the main analysis, results were similar when re-analysing the data without studies that were classified as being at high risk of bias in at least one of the four domains (Table 16). Overall, based on all diagnostic tests and target conditions, the median difference compared to the results from the main analysis was for sensitivity −0.6% (Interquartile range (IQR) −2.1% to 0.7%) and for specificity −0.2% (IQR −0.9% to 0.4%). Excluding studies with high risk of bias had a significant effect (> 10% difference) in 7% (4/54) of the sensitivity or specificity estimates. Notable changes were a decrease in estimated sensitivity of EP for animus from 80% to 63%, of MRI for enterocele from 84% to 73% and for anismus from 86% to 76%. The estimated specificity of MRI for PFD increased from 79% to 91%. Excluding studies at high risk of bias did not notably change the probabilities of the index tests (Table 17), so the conclusions from the main analysis about suitability as a SpIN triage test for rectocele, enterocele, intussusception and PFD, and as a SnOUT triage test for anismus for all index tests, remain valid.

2. Excluding studies with concerns about applicability (women with symptoms of general pelvic floor dysfunction and male participants)

Compared to the main analysis, results were similar when we re-analysed the data without studies including women with symptoms of general pelvic floor dysfunction or male participants (Table 18). Overall, based on all diagnostic tests and target conditions, the median difference compared to the results from the main analysis was for sensitivity −0.7% (IQR −2.6% to 0.1%) and for specificity −0.1% (IQR −1.8% to 0.6%). Concerns about applicability had a significant effect (> 10% difference) in 2% (1/56) of the sensitivity or specificity estimates. The estimated specificity of EVUS for rectocele increased from 76% to 88%. Other notable changes were the decrease in estimated sensitivity of rectocele for TPUS and EVUS with 9%, the increase in sensitivity of EVUS for enterocele and of EP for anismus with 9%. Excluding studies with concerns about applicability did not notably change the probabilities of the index tests for rectocele, intussusception and PFD (Table 19). In women with ODS, EVUS and DAE would no longer be suitable as SpIN triage tests for enterocele, and TPUS, EVUS and EDF would no longer be suitable as SnOUT triage tests for anismus.

3. Excluding studies published before 2010

Compared to the main analysis, results were similar when we re-analysed the data without studies that were published before 2010 (Table 20). Overall, based on all diagnostic tests and target conditions, the median difference compared to the results from the main analysis was for sensitivity −0.7% (IQR −4.5% to 1.1%) and for specificity −0.1% (IQR −1.6% to 0.2%). Excluding older publications had a significant effect (> 10% difference) in 4% (2/56) of the sensitivity or specificity estimates. The estimated sensitivity of EP for anismus decreased from 80% to 67% and of DAE for PFD from 92% to 81%. Other notable changes were an increase in sensitivity of TPUS for rectocele from 88% to 93%, a decrease in sensitivity of MRI for enterocele from 85% to 79% and for anismus from 86% to 80%. Specificity of MRI for rectocele decreased from 90% to 83%. Excluding studies that were published before 2010 did not notably change the probabilities of the index tests (Table 21), except for EVUS for rectocele and EDF for intussusception that would no longer remain suitable as SpIN triage tests.

4. Excluding studies with high risk of bias, studies with concerns about applicability and studies published before 2010

Compared to the main analysis, results were similar when we re-analysed the data without studies that could reduce the overall quality of the evidence (Table 22). Overall, based on all diagnostic tests and target conditions, the median difference compared to the results from the main analysis was for sensitivity −2.6% (IQR −7.1% to −0.5%) and for specificity −1.5% (IQR −3.9% to −0.2%). Excluding studies that reduced the overall quality of evidence had a significant effect (> 10% difference) in 9% (5/54) of the sensitivity or specificity estimates. The estimated sensitivity of TPUS for anismus decreased from 92% to 73% and of MRI for PFD decreased from 94% to 83%. The estimated specificity of EVUS for rectocele increased from 76% to 89%, of MRI for PFD from 79% to 94% and for DAE for PFD decreased from 92 to 77%. Excluding studies that could reduce the overall quality of the evidence did not notably change the probabilities of the index tests for sensitivity (Table 23), so the conclusions from the main analysis about the suitability of all index tests as SnOUT triage tests for anismus and the non-suitability for the other conditions remain valid. The probability of specificity remained the same, except for EVUS and DAE for rectocele which would no longer be suitable as a SpIN triage test, and MRI for anismus which would now be suitable as a SpIN triage test.

Overall quality of evidence

For the assessment of the overall quality of the evidence the GRADE criteria of risk of bias, directness, consistency of effect and precision are evaluated for each outcome.

Risk of bias was low or unclear in most studies and potential limitations were unlikely to lower our confidence in the estimated effect (Sensitivity analysis 1). We identified a serious effect in only four of the outcomes, for which we downgraded by one level.

Directness refers to whether the evaluated tests are the exact tests as used in clinical practice and whether the test accuracy is calculated in the population of interest. In the included studies

all tests were performed as part of routine clinical assessment in a population that required diagnostic imaging, even though not all participants had symptoms solely of ODS. Excluding studies assessing women with symptoms of more general pelvic floor dysfunction did not affect most estimates of sensitivity and specificity (Sensitivity analysis 2). A serious effect was only found for one outcome for which we downgraded the quality of evidence by one level for indirectness.

Consistency refers to the homogeneity of the results across studies, e.g. the degree to which results from included studies are similar, with overlapping confidence intervals. In most forest plots we found no relevant heterogeneity. In the few exceptions, we could not explain the heterogeneity by the performance of tests, so we downgraded the quality of evidence by one level for inconsistency. For DAE and EDF we downgraded the quality of evidence by an extra level for inconsistency, because of concerns about generalisability of the consistency, as the estimated sensitivity and specificity were based on studies from the same research group.

The accuracy of EP, MRI and TPUS for rectocele, enterocele and intussusception was found to be precise, considering the CrIs and the large numbers of studies and participants that contributed. We downgraded the overall quality of evidence of MRI for anismus and PFD for imprecision because of the low number of participants. We downgraded the quality of evidence of EVUS, DAE and EDF for all conditions by one level for imprecision because of a small sample size, and in most cases by an extra level because of wide CrIs, if overall quality was not already downgraded for heterogeneity.

DISCUSSION

Summary of main results

The aim of this review was to determine the diagnostic test accuracy of evacuation proctography (EP), dynamic MRI and pelvic floor ultrasound for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome (ODS), and to assess whether MRI or pelvic floor ultrasound could replace EP. We included 39 studies covering 2483 women in the meta-analysis. The summary of main results, including the overall quality of the evidence, is presented in the Summary of findings 1 We provide a visual overview of the pooled estimated sensitivity and specificity in Figure 14 by target condition and Figure 15 by imaging technique.

Figure 14. Summary estimates of diagnostic test accuracy by target condition based on the results of the LCA.

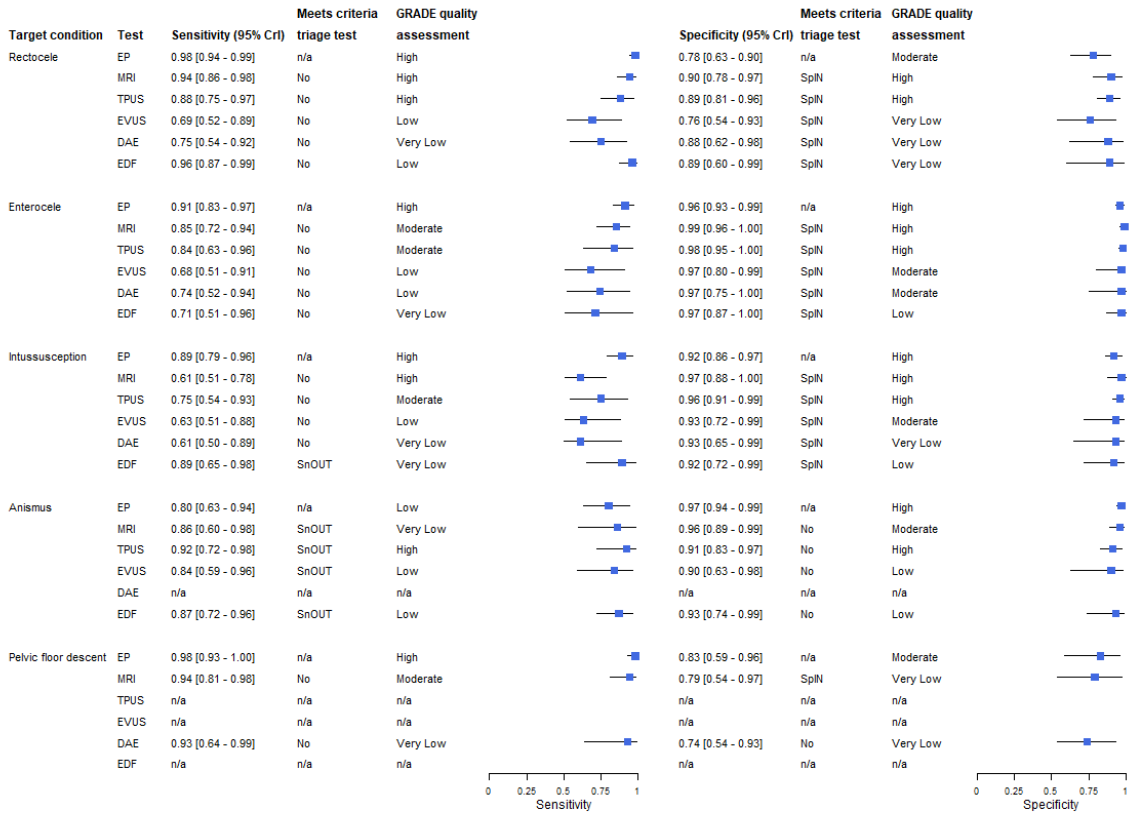
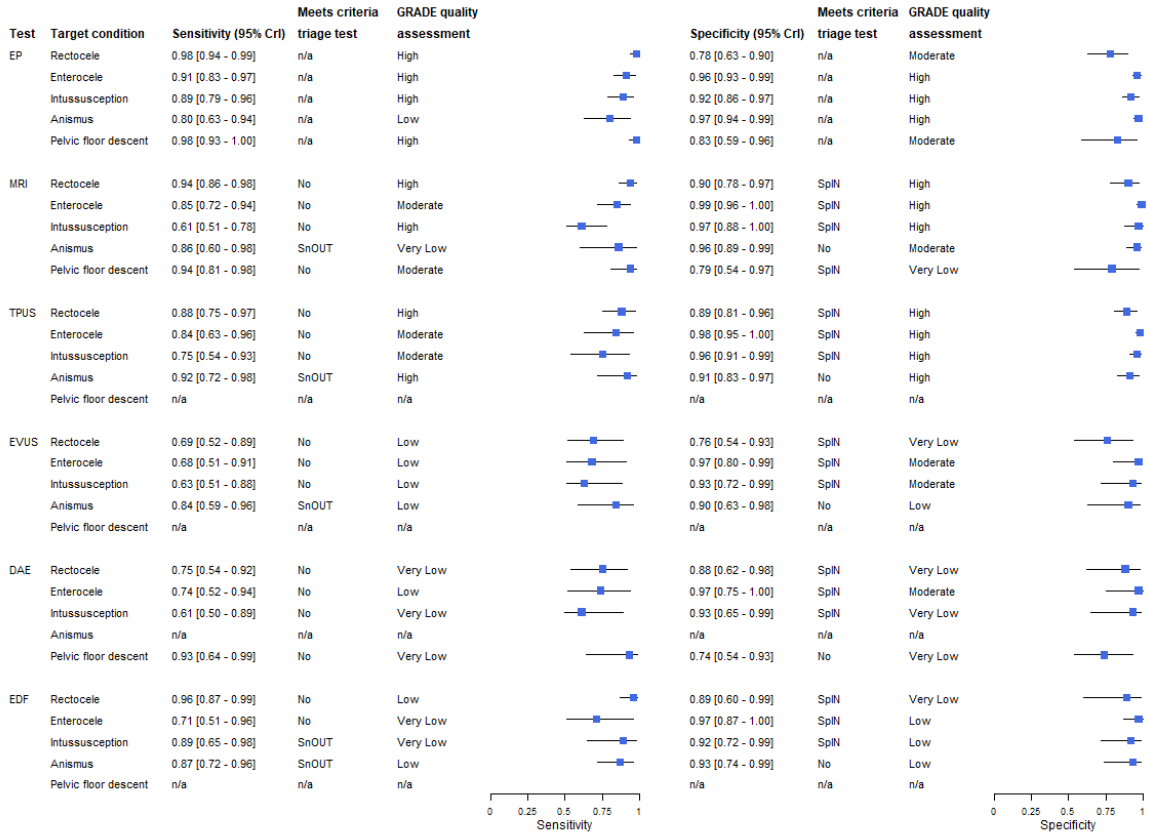


Figure 15. Summary estimates of diagnostic test accuracy by imaging technique based on the results of the LCA.



Diagnostic test accuracy of EP is estimated as follows: for rectocele sensitivity is 98% (CrI 94% to 99%) and specificity 78% (CrI 63% to 90%); for enterocele sensitivity is 91% (CrI 83% to 97%) and specificity 96% (CrI 93% to 99%); for intussusception sensitivity is 89% (CrI 79% to 96%) and specificity 92% (CrI 86% to 97%); for anismus sensitivity is 80% (CrI 63% to 94%) and specificity 97% (CrI 94% to 99%); and for pelvic floor descent sensitivity 98% (CrI 93% to 100%) and specificity 83% (CrI 59% to 96%). High quality of evidence shows that EP has a high sensitivity for diagnosis of rectocele, enterocele, intussusception and pelvic floor descent, and cannot be replaced. Sensitivity of EP for anismus is low, but it cannot be replaced because of its high specificity.

Diagnostic test accuracy of MRI is estimated as follows: for rectocele sensitivity is 94% (CrI 86% to 98%) and specificity 90% (CrI 79% to 97%), for enterocele sensitivity is 85% (CrI 72% to 94%) and specificity 99% (CrI 96% to 100%); for intussusception sensitivity is 61% (CrI 51% to 78%) and specificity 97% (CrI 88% to 100%); for anismus sensitivity is 86% (CrI 60% to 98%) and specificity 96% (CrI 89% to 99%); and for pelvic floor descent sensitivity is 94% (CrI 81% to 98%) and specificity 79% (CrI 54% to 97%). MRI does not meet the criteria to replace EP, but high quality of evidence shows it would be a suitable SpIN triage test for diagnosis of rectocele, enterocele and intussusception. MRI meets the criteria for a SnOUT triage test for anismus and a SpIN triage test for pelvic floor disorder (PFD), but with very low quality of evidence. Heterogeneity analysis shows that sensitivity of MRI performed with an evacuation phase is higher than without evacuation phase for rectocele (94% (CrI 87% to 98%) versus 65% (CrI 52% to 89%)) and enterocele (87% (CrI 74% to 95%) versus 62% (CrI 51% to 88%)), and sensitivity of MRI without evacuation phase is significantly lower than EP (with a probability of 0.001 to 0.013); so MRI should be performed with an evacuation phase.

Diagnostic test accuracy of transperineal ultrasound (TPUS) is estimated as follows: for rectocele sensitivity is 88% (CrI 75% to 97%) and specificity 89% (CrI 81% to 96%); for enterocele sensitivity is 84% (CrI 63% to 96%) and specificity 98% (CrI 95% to 100%), for intussusception sensitivity is 75% (CrI 54% to 93%) and specificity 96% (CrI 91% to 99%), and for anismus sensitivity is 92% (CrI 72% to 98%) and specificity 91% (CrI 83% to 97%). Pelvic floor descent was not assessed using TPUS. TPUS does not meet the criteria to replace EP, but high quality of evidence shows it would be a suitable SpIN triage test for diagnosis of rectocele, enterocele and intussusception and a suitable SnOUT triage test for anismus. Heterogeneity analysis shows that sensitivity of TPUS performed

with rectal contrast is not significantly higher than without rectal contrast for rectocele (92% (CrI 69% to 99%) versus 81% (CrI 58% to 95%)); enterocele (90% (CrI 71% to 99%) versus 67% (CrI 51% to 90%)) and intussusception (90% (CrI 69% to 98%) versus 61% (CrI 51% to 86%)), and is lower than EP (with a probability of 0.125 to 0.529); so rectal contrast is not recommended for clinical use as it is an invasive procedure and EP remains superior.

Diagnostic test accuracy of endovaginal ultrasound (EVUS) is estimated as follows: for rectocele sensitivity is 69% (CrI 52% to 89%) and specificity 76% (CrI 54% to 93%); for enterocele sensitivity is 68% (CrI 51% to 91%) and specificity 97% (CrI 80% to 99%); for intussusception sensitivity is 63% (CrI 51% to 88%) and specificity 93% (CrI 72% to 99%); and for anismus sensitivity is 84% (CrI 59% to 96%) and specificity 90% (CrI 63% to 98%). Pelvic floor descent was not assessed using EVUS. EVUS does not meet the criteria to replace EP. EVUS meets the criteria for a SpIN triage test for diagnosis of rectocele, enterocele and intussusception, and for a SnOUT triage test for anismus, but with moderate to very low quality of evidence.

Diagnostic test accuracy of dynamic anal endosonography (DAE) is estimated as follows: for rectocele sensitivity is 75% (CrI 54% to 92%) and specificity 88% (CrI 62% to 98%); for enterocele sensitivity is 74% (CrI 52% to 94%) and specificity 97% (CrI 75% to 100%); for intussusception sensitivity is 61% (CrI 50% to 89%) and specificity 93% (CrI 65% to 99%), and for pelvic floor descent sensitivity is 93% (CrI 64% to 99%) and specificity 74% (CrI 54% to 93%). Anismus was not assessed using DAE. DAE does not meet the criteria to replace EP. DAE meets the criteria for a SpIN triage test for diagnosis of rectocele, enterocele and intussusception, but with moderate to very low quality of evidence.

Diagnostic test accuracy of echodefaecography (EDF) is estimated as follows: for rectocele sensitivity is 96% (CrI 87% to 99%) and specificity 89% (CrI 60% to 99%); for enterocele sensitivity is 71% (CrI 51% to 96%) and specificity 97% (CrI 87% to 100%); for intussusception sensitivity is 89% (CrI 65% to 98%) and specificity 92% (CrI 72% to 99%); and for anismus sensitivity is 87% (CrI 72% to 96%) and specificity 93% (CrI 74% to 99%). Pelvic floor descent was not assessed using EDF. EDF meets the criteria to replace EP for intussusception, but with a very low quality of evidence. EDF meets the criteria for a SpIN triage test for diagnosis of rectocele and enterocele, and for a SnOUT triage test for anismus, but with low to very low quality of evidence.

Strengths and weaknesses of the review

Strengths and weaknesses of the included studies

The strengths of this review lie in the high number of studies (39) included in the meta-analysis. Most included studies (31/39) prospectively recruited their participants to undergo imaging modalities (Table 3). Studies were performed across the world, and included women with a wide age range. Most of the included studies reported the diagnostic test accuracy of the target conditions in a female population with symptoms of obstructed defaecation, complying with our review question. Most of the studies included a consecutive or random sample of participants, suggesting potentially low selection bias. Most studies performed analysis of the imaging techniques blinded to the comparative test, suggesting potential low detection bias.

Weaknesses of the included studies are that most are single-centre trials (38/39) with a relatively low number of participants; only four studies included more than 100 women (Beer-Gabel 2015; Faggian 2013; Hainsworth 2016; Van Gruting 2017). Twenty per cent (8/39) had a retrospective design. Only three studies investigated the test accuracy of more than one index test (Hainsworth 2016; Van Gruting 2017; Vitton 2011). The included studies showed a wide range in prevalence of the target conditions, which could for instance be caused by difference in prior testing (selection on results on clinical examination rather than only on symptoms), the different cut-off values (e.g. > 0 mm versus > 30 mm), severity of symptoms (any versus daily symptoms of ODS), and settings (tertiary hospitals with more severe cases than secondary). The major limitation is the different methods of performing the imaging techniques and use of cut-off values, potentially causing heterogeneity of the results. Due to a lack of data and co-occurrence of several test performance settings, the heterogeneity analyses give limited insight into these issues. In most studies the imaging technique was only assessed by one examiner, and the interpretation of the index test could therefore be subjective, causing potential review bias. Moreover, it was not reported how well the participants were able to evacuate; insufficient effort might cause underdiagnosis and affects test accuracy. In 13 studies participants with a wider range of symptoms (pelvic floor dysfunction) were included rather than with the sole symptom of obstructed defaecation (Table 3); sensitivity analysis did not show a change in diagnostic test accuracy. In two studies we were not able to retrieve test accuracy data on women only (Ron 2012; Van Iersel 2017), but as these studies included in total 23 men in a total population of 2581 participants we did not expect a significant effect.

Strengths and weaknesses of the review process

The section Differences between protocol and review shows a list of changes, all of which could be seen as a limitation of the review process.

Search strategy and selection process

A strength is that we used no search filters and applied no language or date restrictions. We searched all major electronic databases, including grey literature and those with an appropriate subject focus. We also handsearched references of included studies and checked Google scholar to avoid missing any potential eligible articles. Two review authors independently performed title and abstract screening and full-text evaluation for eligibility, with substantial agreement between them. Studies including men and women, as well as symptomatic and asymptomatic participants were all included, but we requested test accuracy data for symptomatic women only, to minimise selection bias. We requested test accuracy data if not reported, to enable inclusion of these studies.

One limitation is that not all authors provided the requested test accuracy data. Not all studies reporting in a language other than English could be translated and these authors did not all provide test accuracy data on request. This suggests that more studies could have been included in the meta-analysis. Although reporting bias could have occurred, it remains to be established if reporting bias has an impact on the results, as there is currently no reliable method of examining reporting bias for DTA reviews.

Quality assessment and data extraction

A strength is that two review authors performed the data extraction and assessment of methodological quality, with good agreement. We contacted authors of studies from which not all necessary data could be extracted, to provide additional information, even if there was only one item missing. Most of them replied, allowing for minimal missing data. We could include abstracts in the meta-analysis because authors provided information, which reduces publication bias.

A limitation is that not all information was available from the published reports, especially the older ones, as results were not reported using the recently-developed STARD checklist. Although we made extensive efforts to retrieve additional information, not all authors were able to provide us with the necessary information. Another limitation is that the risk of bias was unclear in five studies in three or four domains, and we did not exclude these studies in the sensitivity analysis.

Statistical analysis

The major strength is that we used a Bayesian latent class analysis (LCA) for the meta-analysis, such that the diagnostic test accuracy of all tests could be estimated in one large hierarchical model. Moreover, this model does not require a reference standard, and it provides pooled estimates as well as estimates by study for all relevant diagnostic test accuracy parameters. Furthermore, the Bayesian approach to the LCA makes it possible to estimate the probabilities that the accuracy of one test is equal to or better than the accuracy of a second test. A high number of studies examined the diagnostic test accuracy of EP, MRI and TPUS, enabling the performance of the statistical analysis including heterogeneity and sensitivity analyses, leading to robust results. The LCA model works best when there are studies with more than two tests per participant: there were three large studies with more than two tests. However, the number of included studies for the diagnostic test accuracy of EVUS, DAE, and EDF was low, and we therefore rated their results as low-quality evidence.

As indicated by the secondary objective of the protocol, we aimed to assess the accuracy of each test at prespecified thresholds. However, this was not possible due to the wide range and different definitions of cut-off values used in the included studies. Hence test accuracy estimates were provided for a wide range of thresholds (e.g. rectocele for any cut-off value, ranging from > 0 cm to > 3 cm), which may not correspond to the cut-off values adopted in clinical practice. The heterogeneity evident in the forest plots could be caused by variation in cut-off values between studies, but this is unlikely as no reverse trends in estimates of sensitivities compared to specificities were apparent. Not all intended variations in test performance could be evaluated in the assessment of heterogeneity, because of insufficient data and dependencies between settings. Furthermore, even though EP can be performed with administration of contrast to one or more compartments (rectum, small bowel, vagina or bladder) we did not account for this in the meta-analysis. We acknowledge that the use of multi-compartmental contrast might increase test accuracy and hence could be a potential source of heterogeneity.

Previous research

To date no systematic review or meta-analysis has been performed including all types of imaging techniques for ODS. Ramage 2017 assessed MRI versus either clinical examination or fluoroscopic techniques or both within the same cohort of participants. They compared the detection rates (true positives) and missed rates (false negatives), which reflects estimation of sensitivity, of EP and MRI to avoid needing to use a reference standard. Compared to EP, MRI had a lower detection rate and a higher miss rate for rectoceles, intussusception and perineal descent. This is in line with our results, as sensitivity of MRI for rectocele, enterocele and intussusception was lower compared to EP. The authors did not examine the specificity of MRI and EP. Grossi 2018 assessed differences in diagnostic rates between EP and MRI in participants with constipation as a secondary objective of their meta-analysis. Results were based on pooled prevalence, thereby avoiding the use of a reference standard. They included only studies with more than 40 participants, resulting in analysis of only five studies. EP was superior to MRI for detection of intussusception, which is similar to our findings.

Applicability of findings to the review question

Review question: What is the diagnostic test accuracy of EP, MRI and pelvic floor ultrasound for the diagnosis of posterior pelvic floor disorders in women with obstructed defaecation syndrome?

Patient selection

Narrow inclusion criteria by gender and setting in which the participants are assessed are set. Most included studies therefore report test accuracy for women in secondary or tertiary gynaecology or colorectal surgery outpatient clinics. Studies reporting primarily on men are excluded. Hence test accuracy results of this meta-analysis are only applicable to women seeking help for their ODS symptoms in secondary and tertiary hospitals (women presenting to gynaecologist or colorectal surgeon) and are not applicable to the general population (women presenting to general practitioners) or to male patients.

Wide inclusion criteria by symptoms are defined. Studies included in this review report on women with a variety of symptoms including faecal incontinence, pelvic organ prolapse and more specifically, ODS. Studies that report on asymptomatic women are excluded. Because of the wide range of symptoms, sensitivity analysis could be performed reassessing test accuracy in women with specific symptoms of ODS (27 studies). As we found no difference in test accuracy between the main analysis and the sensitivity analysis, the results of test accuracy in this meta-analysis are applicable to women presenting with general symptoms of pelvic floor dysfunction, irrespective of having specific symptoms of ODS. The test accuracy results of this meta-analysis are not applicable to asymptomatic women.

Women recruited in the included studies had a wide age range (20 to 95 years) and only one study reported only on postmenopausal women (Gufler 2004). Most included studies recruited women with a wide range of variables, e.g. both nulliparous and multiparous, women with various body mass index values, women with or without previous hysterectomy, pelvic floor reconstructive surgery, anorectal or abdominal surgery. Only three studies excluded women with previous pelvic floor surgery (Faucheron 2014; Grasso 2007; Vitton 2011). This suggests that the test accuracy results are generalisable to all women, regardless of age, parity, body mass index and previous surgery. It remains unclear whether results are applicable to mixed-race or mainly white, as in most included studies ethnicity was not reported.

Index test(s)

EP investigations are in most included studies performed in the upright position with the use of rectal contrast and evacuation phase, so test accuracy results should be applicable to clinical practice.

Studies included in this review used either 1T or 1.5T magnet dynamic MRI when performed in the supine position and 0.25T or 0.5T when performed in the upright position. With advances in technology, the quality of the imaging techniques would increase, thereby potentially improving test accuracy results. It is likely that more recently introduced MRI and ultrasound scanners would have better test accuracy than that reported in this review, but sensitivity analysis did not show any increment in accuracy, so test accuracy results should be applicable to all types of MRI scanners.

Although specific types of ultrasound tests are predefined (e.g. transperineal and endovaginal ultrasound), different ultrasound machines and probes are used across the literature. The included studies predominantly used ultrasound scanners of BK Medical and General Electrics, but the use of Siemens, Hitachi and HDI has also been reported. As a variety of ultrasound scanners were included in this review, test accuracy results could be applicable to any type of ultrasound machine, although no heterogeneity analysis has been performed.

The level of experience of operators assessing the index test is not always reported in the included studies; it is either not mentioned (n = 11) or only described as experienced without quantification (n = 16). Only less than a third (n = 12) report the level of experience specified by time (months/years) or in numbers of scans examined before beginning the study. As imaging analysis is subjective and performance highly dependent on training, we assume that test accuracy results of this review are only applicable when imaging techniques are analysed by experienced operators, without being able to qualify this level of experience. In most included studies imaging analysis was performed by a single operator (n = 33). Only in seven studies were images examined by two observers, after which diagnosis was obtained through consensus in case of discrepancies. In clinical practice it is more likely that images are examined by only one operator, as this is more cost-effective. Test accuracy results of index tests presented in this review are therefore applicable to clinical practice where images are assessed by one investigator.

Across the included studies, a variety of test-positive thresholds are used to determine diseased and non-diseased status. Thresholds for diagnosis of target conditions on imaging are based on subjective judgements rather than numeric values, although these thresholds were not always made explicit in the included studies. The specified common threshold therefore still includes a distribution of implicit thresholds. Test accuracy results of this review should be applicable to clinical practice, as it is assumed that this distribution of implicit thresholds is representative of the thresholds used in clinical practice.

AUTHORS' CONCLUSIONS

Implications for practice

In a population of women seeking help for their symptoms of obstructed defaecation, EP remains the best diagnostic imaging technique and cannot be replaced. MRI and TPUS could be used as a triage test, as a positive test confirms diagnosis of rectocele, enterocele and intussusception, and a negative test rules out anismus. MRI should be performed with an evacuation phase. TPUS should not be performed with rectal contrast. Quality of evidence for EVUS, DAE and EDF was too low to support recommendations.

EP remains the best available imaging technique, as it has the highest sensitivity for detection of most of the posterior compartment disorders. However, this review indicates that it should not be called the reference standard, because the specificity of EP was lower compared to the other imaging techniques, with a higher number of false-positives, suggesting the potential of EP in over-diagnosing these conditions.

The harm of misdiagnoses (false-positive and false-negative) is not high for these benign conditions. Women with false-negative results may encounter a delay in treatment, but most women with posterior pelvic floor disorders initially undergo conservative management and therefore a delay by underdiagnosis would not have a major clinical impact on the management of these patients. Women with false-positive results may suffer from unnecessary surgical treatment, but not all of them who are diagnosed with posterior pelvic floor disorders would require surgical intervention. The decision to perform surgery in such women is based not only on imaging findings, but on clinical examination and the severity of symptoms. Most women who require surgical intervention will first receive conservative treatment. If symptoms improve and quality of life is acceptable, no surgical intervention is necessary. As the number of false-negatives and false-positives are not of crucial importance, test accuracy results may range within reasonable limits.

The choice of mode of imaging could now be based on other criteria such as availability, preference, risk assessment and expertise. MRI-defaecography could be preferable to women of child-bearing age, as no radiation is involved. MRI could be performed when there are multi-compartment disorders or in cases of recurrence of symptoms, as it provides a global view of the entire pelvic floor, including supporting structures such as muscles and fascia. MRI-defaecography is not available in every hospital, is more expensive and has contra-indications. MRI could still underdiagnose conditions, especially when the woman is unable to empty her bowels during examination; in these cases, EP should be performed to confirm the diagnosis. MRI with evacuation phase and EP are both invasive investigations with similar patient acceptability (Van Gruting 2017), which suggests that MRI with an evacuation phase is not necessarily more patient-friendly. Ultrasound is less expensive, widely available and more patient-friendly (Perniola 2008; Steensma 2010; Van Gruting 2017). Given that this meta-analysis has shown TPUS to have as high a specificity as EP for most conditions, it could be used as a screening tool in the assessment of women with ODS. TPUS could be performed in women in whom imaging is necessary. When a posterior pelvic floor disorder is found (test positive), it is highly likely that this condition is present and further imaging may not be necessary. However, if conservative therapy has been unsuccessful or if surgery is being contemplated the clinician could choose to perform additional imaging with EP to identify the false-negative cases. Consequently, there could be a reduction in healthcare costs, embarrassment to the woman and length of waiting time for additional imaging.

Implications for research

Low quality of evidence of pelvic floor ultrasound (except from TPUS) requires more well-designed studies, to define its role in the diagnostic pathway of ODS and to enable clinical use:

- Studies to assess test accuracy of DAE and EDF. Both methods use dynamic anorectal ultrasound with the use of rectal contrast and evacuation phase, which potentially does not reduce participant discomfort compared to EP and dynamic MRI. The scanning protocol requires extensive training, as imaging is performed during evacuation of contrast. Studies of DAE and EDF are mainly conducted by the same principal investigator, i.e. Regadas for EDF and Barthet for DAE. These techniques were developed more than a decade ago, but so far no other centres have experience with these ultrasound methods for assessment of ODS. Before these techniques can be implemented in clinical practice, diagnostic test accuracy, reproducibility, generalisability and patients' experience need to be evaluated in a wider setting.
- Studies to assess test accuracy of EVUS. EVUS is a recently-developed and potentially useful imaging technique for the assessment of ODS, especially in combination with other forms of ultrasound (Hainsworth 2016; Van Gruting 2017). Test accuracy has been assessed by different authors in different countries, but more studies to assess test accuracy are necessary to empower results.

Future well-designed diagnostic studies are recommended to assess imaging techniques for their test accuracy for diagnosis of posterior pelvic floor disorders for which this meta-analysis was not able to provide clear results:

- Studies on pelvic floor descent because of a low number of studies. Pelvic floor descent is a more recently identified possible cause of ODS, so not many authors have included this condition in their test accuracy assessment. It has the highest prevalence of all posterior compartment disorders (67%) and should be taken into consideration when assessing women with symptoms of ODS. A few studies of pelvic floor descent on EP and MRI are published, but it has yet to be established how pelvic floor descent should be defined on ultrasound, including establishment of cut-off values.
- Studies on MRI in the upright position. EP is superior to MRI in the supine position with evacuation phase. It remains unknown if MRI in the upright position would have a similar test accuracy compared to EP. Two studies show excellent results (Fiaschetti 2013; Schoenenberger 1998), but more studies are needed to empower these findings. A specially dedicated open MRI scanner is required for these examinations, so these studies may only be conducted in large academic centres. Availability of these open-MRI scanners might influence the likelihood of implementing this method in general clinical practice.
- Studies assessing the effect of different cut-off values on the test accuracy of imaging. Currently a wide range of different classification systems between and within imaging techniques exist. For example, different reference lines could be used for the assessment of pelvic floor disorders on MRI. It has been widely accepted that the pubococcygeal line provides the most accurate measurements. Other lines, such as the midpubic line, which corresponds to hymenal remnant, could also be used. Similarly, on ultrasound, the line parallel to the lower aspect of pubic symphysis is considered as a reference, but this does not correspond to the lines used for EP and MRI. Further studies are therefore required to assess different cut-off values and reference lines to enable standardisation and global implementation. Moreover, the use of uniform cut-off values for all imaging techniques is essential in meta-analyses, to reduce heterogeneity and provide reliable test accuracy data.

Additional research is needed on aspects of tests beyond test accuracy:

- Studies assessing cost effectiveness. MRI is a more expensive imaging technique compared to the relatively cheap EP. It needs to be established whether MRI is more expensive than EP in the long term. It could well be that advanced knowledge of the multi-compartment diagnoses would lead to appropriate surgery being performed at the outset, thereby minimising the risk of recurrence of symptoms or failed surgery, and avoiding multiple interventions. When using relatively low-cost ultrasound for the initial assessment of women with ODS, these women may then not require additional EP or MRI when symptoms improve after conservative treatment for their condition. This could potentially reduce the number of women requiring EP or MRI and consequently reducing healthcare costs.

ACKNOWLEDGEMENTS

CCCG editorial office for help with search strategies and copy editing.
Nandini Dendukuri for her guidance using the LCA approach.
Joris Menten for his help with syntax and model choice.
Authors of included articles for providing additional data on request.

REFERENCES

References to studies included in this review

Barthet 2000 *{published data only}*

Barthet M, Houtin D, Pottier F, Bouvier M, Mambrini P, Picon M, et al. Prospective comparison of endosonography and defecography in the diagnosis of anorectal dynamic disorders and perineal insufficiency. *Gastroenterology* 1998;**114**(4):A718.

* Barthet M, Portier F, Heyries L, Orsoni P, Bouvier M, Houtin D, et al. Dynamic anal endosonography may challenge defecography for assessing dynamic anorectal disorders: results of a prospective pilot study. *Endoscopy* 2000;**32**(4):300-5.

Beer-Gabel 2004 *{published and unpublished data}*

Beer-Gabel M, Teshler M, Schechtman E, Zbar AP. Dynamic transperineal ultrasound vs. defecography in patients with evacuatory difficulty: a pilot study. *International Journal of Colorectal Disease* 2004;**19**(1):60-7.

Beer-Gabel 2008 *{published and unpublished data}*

Beer-Gabel M, Assoulin Y, Amitai M, Bardan E. A comparison of dynamic transperineal ultrasound (DTP-US) with dynamic evacuation proctography (DEP) in the diagnosis of cul de sac hernia (enterocele) in patients with evacuatory dysfunction. *International Journal of Colorectal Disease* 2008;**23**(5):513-9.

Beer-Gabel 2015 *{published and unpublished data}*

Beer-Gabel M, Carter D. Comparison of dynamic transperineal ultrasound and defecography for the evaluation of pelvic floor disorders. *International Journal of Colorectal Disease* 2015;**30**(6):835-41.

Brusciano 2007 *{published and unpublished data}*

Brusciano L, Limongelli P, Pescatori M, Napolitano V, Gagliardi G, Maffettone V, et al. Ultrasonographic patterns in patients with obstructed defaecation. *International Journal of Colorectal Disease* 2007;**22**(8):969-77.

Dellemare 1994 *{published data only (unpublished sought but not used)}*

Delemarre JB, Kruyt RH, Doornbos J, Buyze-Westerweel M, Trimbos JB, Hermans J, et al. Anterior rectocele: assessment with radiographic defecography, dynamic magnetic resonance imaging, and physical examination. *Diseases of the Colon and Rectum* 1994;**37**(3):249-59.

Faggian 2013 *{published data only (unpublished sought but not used)}*

Faggian A, Alabiso ME, Serra N, Pizza NL, Iasiello F, Tecame M, et al. Entero-colpo-defecography vs supine entero-MRI: which one is the best tool in the differentiation of enterocele, elytrocele and edrocele? *Journal of Biological Regulators and Homeostatic Agents* 2013;**27**(3):861-8.

Faucheron 2014 *{published and unpublished data}*

Faucheron JL, Barot S, Collomb D, Hohn N, Anglade D, Dubreuil A. Dynamic cystocolpoproctography is superior to functional pelvic MRI in the diagnosis of posterior pelvic floor disorders: results of a prospective study. *Colorectal Disease* 2014;**16**(7):O240-7. [DOI: 10.1111/codi.12586]

Fiaschetti 2013 *{published data only (unpublished sought but not used)}*

De Luca E, Sileri PP, Franceschilli L, Angelucci G, Arcudi C, Perrone F, et al. Laparoscopic ventral rectopexy for internal rectal prolapse. In: Techniques in Coloproctology. Vol. 17. 2013:142.

* Fiaschetti V, Pastorelli D, Squillaci E, Funel V, Rascioni M, Meschini A, et al. Static and dynamic evaluation of pelvic floor disorders with an open low-field tilting magnet. *Clinical Radiology* 2013;**68**(6):e293-300.

Franceschilli L, Lazzaro S, Angelucci GP, De Luca E, Cadeddu F, Milito G, et al. Laparoscopic ventral mesh rectopexy for internal rectal prolapse. In: Techniques in Coloproctology. Vol. 16. 2012:97-8.

Franceschilli L, Lazzaro S, Fiaschetti V, Funel V, Angelucci G, Patrizi L. Laparoscopic ventral rectopexy for rectal prolapse using biological mesh. *Diseases of the Colon and Rectum* 2011;**54**(5):e32-e33.

Franceschilli L, Varvaras D, Capuano I, Ciangola CI, Giorgi F, Boehm G, et al. Laparoscopic ventral rectopexy using biologic mesh for the treatment of obstructed defaecation syndrome and/or faecal incontinence in patients with internal rectal prolapse: a critical appraisal of the first 100 cases. *Techniques in Coloproctology* 2015;**19**(4):209?19.

Sileri P, Franceschilli L, De Luca E, Arcudi C, Giorgi F, Gaspari AL. Laparoscopic ventral rectopexy using biological cross-linked mesh. *Colorectal Disease* 2012;**14**(Suppl 2):15.

Sileri P, Franceschilli L, Larrazo S, Angelucci GP, Patrizi L, Fiaschetti V, et al. Laparoscopic ventral rectopexy for rectal prolapse using biological mesh. In: Gastroenterology. 2011:S1013.

Sileri P, Franceschilli L, Lazzaro S, De Luca E, Angelucci GP, Gaspari AL. Laparoscopic ventral mesh rectopexy for internal rectal prolapse using biological meshes. *Colorectal Disease* 2011;**13**(Suppl 6):12.

Sileri P, Franceschilli L, Limura E, Gaspari AL, Perrone F. Pelvic organs prolapse surgery: Pelvic organ suspension (POPS), laparoscopic ventral mesh rectopexy or starr? *Surgical Endoscopy* 2014;**28**:S123.

Sileri P, Franceschilli L, de Luca E, Lazzaro S, Angelucci GP, Fiaschetti V, et al. Laparoscopic ventral rectopexy for internal rectal prolapse using biological mesh: postoperative and short-term functional results. *Journal of Gastrointestinal Surgery* 2012;**16**(3):622-8.

Sileri P, Shalaby M, Francheschilli L, Missori G, Quaresima S, Capuano I. Laparoscopic ventral bio-mesh rectopexy for internal and external rectal prolapse. *Gastroenterology* 2016:S1262.

Foti 2013 *{published and unpublished data}*

Foti PV, Farina R, Riva G, Coronella M, Fisichella E, Palmucci S, et al. Pelvic floor imaging: comparison between magnetic resonance imaging and conventional defecography in studying outlet obstruction syndrome. *Radiologia Medica* 2013;**118**(1):23-39. [DOI: 10.1007/s11547-012-0840-8]

Grasso 2007 *{published data only (unpublished sought but not used)}*

Grasso RF, Piciucchi S, Quattrocchi CC, Sammarra M, Ripetti V, Zobel BB. Posterior pelvic floor disorders: a prospective comparison using introital ultrasound and colpocystodefecography. *Ultrasound in Obstetrics & Gynecology* 2007;**30**(1):86-94.

Gufler 1999 *{published and unpublished data}*

Gufler H, Laubenberger J, DeGregorio G, Dohnicht S, Langer M. Pelvic floor descent: dynamic MR imaging using a Half-Fourier RARE Sequence. *Journal of Magnetic Resonance Imaging* 1999;**9**(3):378-83.

Gufler 2004 *{published and unpublished data}*

Gufler H, Ohde A, Grau G, Grossmann A. Colpocystoproctography in the upright and supine positions correlated with dynamic MRI of the pelvic floor. *European Journal of Radiology* 2004;**51**(1):41-7.

Hainsworth 2016 *{published and unpublished data}*

Hainsworth A, Collins E, Solanki D, Griffin N, Schizas A, Williams A. Inter- and intra- rater repeatability of the measurement of rectocoele using transperineal ultrasound and correlation with defaecation proctography. *Colorectal Disease* 2015;**17**:40.

Hainsworth A, Collins E, Solanki D, Griffin N, Schizas AM, Williams AB. The measurement and assessment of rectocoele using ultrasound and proctography. *Gut* 2015;**64**:A377.

Hainsworth A, Morris S, Solanki D, Schizas A, Williams A. Prospective validation of integrated total pelvic floor (transperineal, transvaginal) ultrasound in pelvic floor defaecatory dysfunction by comparison with defaecation proctography. *Neurourology and Urodynamics* 2017;**36**(Supplement 3):S316-8.

* Hainsworth A, Solanki D, Hamad A, Morris S, Schizas A, Williams A. Integrated total pelvic floor ultrasound in pelvic floor defaecatory dysfunction. *Colorectal Disease* 2016;**19**:O54-65.

Hainsworth A, Solanki D, Lyons M, Schizas AM, Williams AB. Total pelvic floor ultrasound in pelvic floor defaecatory dysfunction. In: Gut. Vol. 64. 2015:A376-7.

Hainsworth A, Solanki D, Schizas A, Williams A. Total pelvic floor ultrasound in pelvic floor defaecatory dysfunction. *Neurourology and Urodynamics* 2015;**34**:S122-3.

Hainsworth A, Solanki D, Schizas A, Williams A. Total pelvic floor ultrasound in the colorectal pelvic floor multidisciplinary meeting. *Colorectal Disease* 2015;**17**:40-1.

Hainsworth A. The measurement and assessment of rectocoele using total pelvic floor ultrasound and proctography. *Neurourology and Urodynamics* 2015;**34**:S256-7.

Hainsworth A. Total pelvic floor ultrasound and pelvic floor defaecatory dysfunction. *Colorectal Disease* 2015;**17**:19.

Hainsworth AJ, Solanki D, Morris SJ, Schizas AMP, Williams AB. The prospective validation of integrated total pelvic floor (Transperineal, Transvaginal) ultrasound in pelvic floor defaecatory dysfunction. *Colorectal Disease* 2017;**19**(Supplement 4):47.

Halligan 1996 *{published data only}*

Halligan S1, Northover J, Bartram CI. Vaginal endosonography to diagnose enterocoele. *British Journal of Radiology* 1996;**69**(827):996-9.

Healy 1997 *{published data only (unpublished sought but not used)}*

Healy JC, Halligan S, Reznek RH, Watson S, Bartram CI, Phillips R, et al. Dynamic MR imaging compared with evacuation proctography when evaluating anorectal configuration and pelvic floor movement. *American Journal of Roentgenology* 1997;**169**(3):775-9.

Karaus 2000 *{published data only}*

Karaus M, Neuhaus P, Wiedenmann TB. Diagnosis of enteroceles by dynamic anorectal endosonography. *Diseases of the Colon and Rectum* 2000;**43**(12):1683-8.

Kelvin 2000 *{published data only}*

Kelvin FM, Maglinte DD, Hale DS, Benson JT. Female pelvic organ prolapse: a comparison of triphasic dynamic MR imaging and triphasic fluoroscopic cysto-colpo-proctography. *American Journal of Roentgenology* 2000;**174**(1):81-8.

Lienemann 1997 *{published and unpublished data}*

Anthuber C, Baron A, Lienemann A. Dynamic magnetic resonance colpocystorectography for diagnosis of pelvic floor descent and genital prolapse [Die dynamische Magnetresonanzkolpozystorektographie bei der diagnostik von deszensus und prolaps genitalis]. *Gynaecologe* 1996;**29**:620-3.

* Lienemann A, Anthuber C, Baron A, Kohz P, Reiser M. Dynamic MR colpocystorectography assessing pelvic-floor descent. *European Radiology* 1997;**7**(8):1309-17.

Lienemann A, Anthuber CJ, Baron A, Reiser M. Dynamic MR colpocystorectography. A new methods for evaluating pelvic floor descent and genital prolapse [Dynamische MR-Kolpozystorektographie ein neues verfahren zur beurteilung von deszensus und prolaps genitalis]. *Aktuelle Radiologie* 1996;**6**(4):182-6.

Lienemann 2000 *{published and unpublished data}*

Lienemann A, Anthuber C, Baron A, Reiser M. Diagnosing enteroceles using dynamic magnetic resonance imaging. *Diseases of the Colon and Rectum* 2000;**43**(2):205-13.

Martellucci 2011 *{published and unpublished data}*

Martellucci J, Naldini G, Ricchiuti A, Menconi C, Romani N, Rossi M. Dynamic transperineal ultrasound in patients with obstructed defecation. *Colorectal Disease* 2009;**11**:28.

* Martellucci J, Naldini G. Clinical relevance of transperineal ultrasound compared with evacuation proctography for the evaluation of patients with obstructed defaecation. *Colorectal Disease* 2011;**13**(10):1167-72.

Martin 2017 *{published and unpublished data}*

Martin GM, García-Armengol F, Vila JVR, Sanjuán VM, Pérez MM, Espi-Macias A, et al. Prospective comparative study between dynamic pelvic magnetic resonance and videodefecography in assessing obstructive defecation syndrome: Preliminary results. *Colorectal Disease* 2014;**16**(Suppl. 3):40.

* Martin-Martin GP, Garcia-Armengol J, Roig-Vila JV, Espi-Macias A, Martinez-Sanjuan V, Minguez-Perez M, et al. Magnetic resonance defecography versus videodefecography in the study of obstructed defecation syndrome: Is videodefecography still the test of choice after 50 years? *Techniques in Coloproctology* 2017;**21**(10):795-802.

Matsuoka 2000 *{published and unpublished data}*

* Matsuoka H, Desai MB, Wexner SD, Adami C, Mavrantonis C, Nogueras JJ, et al. A pilot assessment of whether external coil MRI is useful to assess evacuatory disorders. *International Journal of Colorectal Disease* 2000;**15**(2):91-5.

Matsuoka H, Wexner SD, Desai MB, Nakamura T, Nogueras JJ, Weiss EG, et al. A comparison between dynamic magnetic resonance imaging and videoproctography in patients with constipation. *Diseases of the Colon and Rectum* 2001;**44**(4):571-6.

Miravalle 2016 *{published and unpublished data}*

Miravalle OR, Farina PA, Muñoz JP, Arias JH, Bruzzi MS, Gualdrini U, et al. Comparison between defecography and echodefecography in the assessment of anorectal dysfunction in patients with obstructed defecation. *Gastroenterology* 2016; (Suppl 1):S301.

Murad-Regadas 2008 *{published and unpublished data}*

Murad-Regadas SM, Regadas FS, Rodrigues LV, Silva FR, Soares FA, Escalante RD. A novel three-dimensional dynamic anorectal ultrasonography technique (echodefecography) to assess obstructed defecation, a comparison with defecography. *Surgical Endoscopy* 2008;**22**(4):974-9.

Murad-Regadas 2011 *{published and unpublished data}*

Murad-Regadas S, Regadas F, Soares G, Rodrigues L, Buchen G, Kenmoti V. A novel dynamic 3-D anorectal ultrasound technique to assess perineal descent compared with conventional defecography. *Diseases of the Colon and Rectum* 2010;**53**(4):588-9.

* Murad-Regadas SM, Dos Santos D, Soares G, Regadas FS, Rodrigues LV, Buchen G, et al. A novel three-dimensional dynamic anorectal ultrasonography technique for the assessment of perineal descent, compared with defaecography. *Colorectal Disease* 2011;**14**(6):740-7.

Perniola 2008 *{published and unpublished data}*

Perniola G, Shek C, Chong CC, Chew S, Cartmill J, Dietz HP. Defecation proctography and translabial ultrasound in the investigation of defecatory disorders. *Ultrasound in Obstetrics & Gynecology* 2008;**31**(5):567-71.

Pilkington 2012 *{published and unpublished data}*

Pilkington SA, Nugent KP, Brennan J, Harris S, Thomas C, Tarver D. Barium proctography vs dynamic magnetic resonance proctography for pelvic floor disorders. *Colorectal Disease* 2010;**12**:22.

* Pilkington SA, Nugent KP, Brenner J, Harris S, Clarke A, Lamparelli M, et al. Barium proctography vs magnetic resonance proctography for pelvic floor disorders: a comparative study. *Colorectal Disease* 2012;**14**(10):1224-30.

Poncelet 2017 *{published data only}*

* Poncelet E, Rock A, Quinton JF, Cosson M, Ramdane N, Nicolas L, Feldmann A, Salleron J. Dynamic MR defecography of the posterior compartment: Comparison with conventional X-ray defecography. *Diagnostic and Interventional Imaging* 2017;**98**:327-32.

Regadas 2011 *{published data only}*

Patel CB, Ragupathi M, Regadas FS, Sardinas C, Jorge JM, Murad-Regadas SM, et al. Assessment of concordance rate between echodefecography and conventional defecography in patients with obstructed defecation syndrome: an international multicenter trial. *Gastrointestinal Endoscopy* 2010;**71**:AB131.

Regadas F, Haas E, Abbas M, Jorge J, Habr-Gama A, Sands D, et al. Prospective multicenter trial comparing echodefecography with defecography in the assessment of anorectal dysfunctions in patients with obstructed defecation. *Diseases of the Colon and Rectum* 2010;**53**(4):534.

* Regadas FS, Haas EM, Abbas MA, Marcio Jorge J, Habr-Gama A, Sands D, et al. Prospective multicenter trial comparing echodefecography with defecography in the assessment of anorectal dysfunction in patients with obstructed defecation. *Diseases of the Colon and Rectum* 2011;**54**(6):686-92.

Ron 2012 *{published and unpublished data}*

Ron Y, Zelber Sagi S, Dekel R, Tiomny E, Santo E, Halpern Z, et al. Comparison of dynamic transperineal ultrasound with evacuation proctography for the evaluation of patients with defecation disorders. *Gastroenterology* 2012;**142**(5):S901.

Steensma 2010 *{published and unpublished data}*

* Steensma AB, Oom DM, Burger CW, Schouten WR. Assessment of posterior compartment prolapse: a comparison of evacuation proctography and 3D transperineal ultrasound. *Colorectal Disease* 2010;**12**(6):533-9.

Steensma AB, Oom DMJ, Burger CW, Schouten WR. Comparison of defecography and 3d/4d translabial ultrasound in patients with pelvic organ prolapse and/or evacuation disorders. *International Urogynecology Journal* 2007;**18**(Suppl 1):S81.

Vanbeckevoort 1999 *{published data only}*

Vanbeckevoort D, Van Hoe L, Oyen R, Ponette E, De Ridder D, Deprest J. Pelvic floor descent in females: comparative study of colpocystodefecography and dynamic fast MR imaging. *Journal of Magnetic Resonance Imaging* 1999;**9**(3):373-7.

Van Gruting 2017 *{published and unpublished data}*

Van Gruting IM, Kluivers K, Sultan AH, De Bin R, Stankiewicz A, Blake H, et al. Does 4D transperineal ultrasound have additional value over2D transperineal ultrasound for diagnosing posterior pelvic floor disorders in women with obstructed defecation syndrome? *Ultrasound in Obstetrics & Gynecology* 2018;**52**(6):784?91.

Van Gruting IM, Stankiewicz A, Kluivers K, Blake H, Thakar R, Sultan AH. Accuracy of evacuation proctogram, MRI and ultrasound for detecting posterior compartment disorders. *International Urogynecology Journal* 2015;**26 (Supplement 1)**:S69-70.

* Van Gruting IM, Stankiewicz A, Kluivers K, De Bin R, Blake H, Sultan AH, et al. Accuracy of four imaging techniques for diagnosis of posterior pelvic floor disorders. *Obstetrics and Gynecology* 2017;**130**(5):1017-24.

Van Iersel 2017 *{published data only}*

Van Iersel JJ, Formijne Jonkers HA, Verheijen PM, Broeders IA, Heggelman BG, Fiitterer JJ, et al. Prospective comparison of dynamic MR defecography with rectal evacuation and conventional defecography for prolapse of the posterior compartment. *Surgical Endoscopy and Other Interventional Techniques* 2017;**31**(2 Supplement 1):S356.

* Van Iersel JJ, Formijne Jonkers HA, Verheijen PM, Broeders IA, Heggelman BG, Sreetharan V, et al. Comparison of dynamic magnetic resonance defaecography with rectal contrast and conventional defaecography for posterior pelvic floor compartment prolapse. *Colorectal Disease* 2017;**19**(1):O46-53.

Vitton 2011 *{published data only}*

Vitton V, Vignally P, Barthet M, Cohen V, Durieux O, Bouvier M, et al. Dynamic anal endosonography and MRI defecography in diagnosis of pelvic floor disorders: comparison with conventional defecography. *Diseases of the Colon and Rectum* 2011;**54**(11):1398-404.

Weemhoff 2013 *{published and unpublished data}*

Weemhoff M, Kluivers KB, Govaert B, Evers JL, Kessels AG, Baeten CG. Transperineal ultrasound compared to evacuation proctography for diagnosing enteroceles and intussusceptions. *International Journal of Colorectal Disease* 2011;**28**(3):359-63.

Zafar 2012 *{published and unpublished data}*

Zafar A, Chapman M. Evacuation proctography and magnetic resonance defaecography: a retrospective study. *Pelvi-Perineology* 2012;**31**(4):107-8.

Zafar 2017 *{published and unpublished data}*

Feretis M. Comparative study of magnetic resonance defaecography and evacuation proctography in evaluation of pelvic floor dysfunction. WHO ICTRP (International Clinical Trials Registry Platform) at apps.who.int/trialsearch/ (first received 18 June 2012). [ISRCTN73466412]

* Zafar A, Seretis C, Feretis M, Karandikar S, Williams SC, Goldstein M, et al. Comparative study of magnetic resonance defaecography and evacuation proctography in the evaluation of obstructed defaecation. *Colorectal Disease* 2017;**19**(6):O204-9.

References to studies excluded from this review

Beer-Gabel 2002 *{published data only (unpublished sought but not used)}*

Beer-Gabel M, Teshler M, Barzilai N, Lurie Y, Malnick S, Bass D, et al. Dynamic transperineal ultrasound in the diagnosis of pelvic floor disorders. *Diseases of the Colon and Rectum* 2002;**45**(2):239-48.

Beer-Gabel 2010 *{published data only (unpublished sought but not used)}*

Beer-Gabel M, Iskhakov A, Khaikin M, Bar Meir S. Obstructive defecation syndrome: Is it only a rectal disorder? *Techniques in Coloproctology* 2010;**14**:85.

Beer-Gabel 2011 *{published data only (unpublished sought but not used)}*

Beer-Gabel M, Iskhakov A, Carter D, Bardan E, Zbar AP. Constipation-predominant irritable bowel syndrome and obstructed defaecation syndrome: Is there diagnostic overlap? *Techniques in Coloproctology* 2011;**15**:228.

Bot-Robin 2011 *{published data only (unpublished sought but not used)}*

Bot-Robin V, Drain A, Lucot JP, Poncelet E, Quinton JF, Cosson M. Assessment of vaginal rectopexy in the concomitant treatment of rectal and genital prolapse with mesh [Faisabilité du traitement concomitant du prolapsus rectalet génital par prothèse par voie vaginale avec rectopexie]. *Pelvi-Perineologie* 2011;**6**:166-73.

Bussen 2003 *{published data only (unpublished sought but not used)}*

Bussen D, Kenn W, Stoffels J, Moll R, Sailer M. Comparative study of dynamic MR-imaging and evacuation proctography in patients with pelvic floor syndrome [Vergleich der dynamischen MR-Defäkographieund konventionellen Röntgendefäkographie inder Diagnostik der Beckenbodeninsuffizienz]. *Coloproctology* 2003;**25**(6):301-7.

Cappabianca 2011 *{published data only (unpublished sought but not used)}*

Cappabianca S, Reginelli A, Iacobellis F, Granata V, Urciuoli L, Alabiso ME, et al. Dynamic MRI defecography vs. entero-colpo-cystodefecography in the evaluation of midline pelvic floor hernias in female pelvic floor disorders. *International Journal of Colorectal Disease* 2011;**26**(9):1191-6.

Cerdán 2011 *{published data only (unpublished sought but not used)}*

Cerdán C, Vigara M Ortega M, Jimenez F, Cerdán J. Enterocele in the elderly: Important obstructive defecation etiopathogenic

factor. Surgical treatment results. *European Geriatric Medicine* 2011;**2S**:S187.

Chatoor 2007 *{published data only (unpublished sought but not used)}*

Chatoor DR, Emmanuel AV, Elneil S, Cohen R, Windsor A, Osborne J, et al. What does MR proctography add in comparison to fluoroscopic proctography in patients iwth evacuation difficulty? *Gut* 2007;**56**:A47.

* Chatoor DR, Emmanuel AV, Elneil S, Cohen R, Windsor A, Osborne J, et al. What does MR proctography add in comparison to fluoroscopic proctography in patients with evacuation difficulty? *Gastroenterology* 2007;**132**(4):A355.

Chung 2003 *{published data only (unpublished sought but not used)}*

Chung EJ, Myang SJ, Kim AY, Kim TH, Lee JH, Yang SK, et al. Utility of dynamic MR proctography in obstructive defaecation. *Gastroenterology* 2003;**124**(4):A684.

Dekel 2015 *{published data only (unpublished sought but not used)}*

Dekel R, Hod K and Ron Y. The value of a balloon expulsion test in the workup of obstructed defecation. *Gastroenterology* 2015;**1**:S305.

Deval 2003 *{published data only (unpublished sought but not used)}*

Deval B, Vulierme MP, Poilpot S, Menu Y, Levardon. Imagerie du prolapsus génito-urinaire [Imaging pelvic floor prolapse]. *Journal de Gynecologie, Obstetrique et Biologie de la Reproduction* 2003;**32**(1):22-9.

Dvorkin 2004 *{published data only (unpublished sought but not used)}*

Dvorkin LS, Hetzer F, Scott SM, Williams NS, Gedroyc W, Lunniss PJ. Open-magnet MR defaecography compared with evacuation proctography in the diagnosis and management of patients with rectal intussusception. *Colorectal Disease* 2004;**6**(1):45-53.

Ferrari 2019 *{published data only (unpublished sought but not used)}*

Ferrari L, Cuinas K, Schizas A, Darakhshan A, Williams A. Overview of patients with primary faecal incontinence symptoms: conservative treatment and surgical intervention based on symptoms' stratification. *Neurourology and Urodynamics* 2019;**38 (Supplement 3)**:S108-10.

Fletcher 2003 *{published data only}*

* Fletcher JG, Busse RF, Riederer SJ, Hough D, Gluecker T, Harper CM, et al. Magnetic resonance imaging of anatomic and dynamic defects of the pelvic floor in defecatoryDisorders. *American Journal of Gastroenterology* 2003;**98**(2):399-411.

Fletcher JG, Klingele CJ, Gebhart JB, Helwig P, Riederer SJ, Bharucha AE, et al. Mr proctography for identifying pelvic floor dysfunction in obstructed defecation. *Gastroenterology* 2002;**122**:A447.

Goffredo 2010 *{published data only}*

Goffredo F, Zirizzotti G, Gervasi C, Ciarletti S, Gioieni A, Bonfà M, et al. Combined role of anorectal manometry and transperineal ultrasonography in the diagnosis of outlet constipation. *Digestive and Liver Disease* 2010;**42S**:S179.

Groenendijk 2009 *{published data only (unpublished sought but not used)}*

Groenendijk AG, Birnie E, De Blok S, Adriaanse AH, Ankum WM, Roovers JP, et al. Clinical-decision taking in primary pelvic organ prolapse; the effects of diagnostic tests on treatmentselection in comparison with a consensus meeting. *International Urogynecology Journal and Pelvic Floor Dysfunction* 2009;**20**:711-9.

Healy 1998 *{published data only (unpublished sought but not used)}*

Healy JC, Halligan S, Reznek RH, Watson S, Bartram CI, Kamm MA, et al. Magnetic resonance imaging of the pelvic floor in patients with obstructed defaecation. *British Journal of Surgery* 1998;**84**(11):1555-8.

Imanova 2017 *{published data only}*

Imanova SS, Sultanova MC. The importance of the dynamic X-ray defecography diagnosis of anorectum diseases. *Azerbaijan Medical Journal* 2017;**1**:114-20.

Kaufman 2001 *{published data only (unpublished sought but not used)}*

Kaufman HS, Buller JL, Thompson JR, Pannu HK, DeMeester SL, Genadry RR, et al. Dynamic pelvic magnetic resonance imaging and cystocolpoproctography alter surgical management of pelvic floor disorders. *Diseases of the Colon and Rectum* 2001;**44**(11):1575-83.

Kawata 2010 *{published data only (unpublished sought but not used)}*

Kawata M, Shah N, Woodfield C, Murphy B, Husain S, Pricolo V. Posterior pelvic floor dysfunction: a comparison of defecography, MRI, and anorectal physiology. *Diseases of the Colon and Rectum* 2010;**53**(4):589.

Köhler 2012 *{published data only (unpublished sought but not used)}*

Köhler K, Stelzner S, Hellmich G, Lehmann D, Jackisch T, Fankhänel B, et al. Results in the long-term course after stapled transanal rectal resection (STARR). *Coloproctology* 2013;**35**(3):169-76.

* Köhler K, Stelzner S, Hellmich G, Lehmann D, Jackisch T, Fankhänel B, et al. Results in the long-term course after stapled transanal rectal resection (STARR). *Langenbecks Archives of Surgery* 2012;**397**(5):771-8.

Mege 2013 *{published and unpublished data}*

Mege D, Ouaissi M, Pirro N, Sastre B, Sielezneff I. Predictive factors for long-term symptomatic failure following elytrocele surgical correction. *European Surgical Research* 2013;**50**:202. [DOI: 10.1159/000351963]

Ortega 2011 *{published data only (unpublished sought but not used)}*

Ortega M, Jimenez F, Sanz G, Esteban F, Garcia M, Cerdán J. Enterocoele: results of surgical treatment. In: Colorectal Disease. Vol. 13. 2011:49.

Otto 2011 *{published data only (unpublished sought but not used)}*

Otto SD, Oesterheld A, Ritz JP, Gröne J, Wolf KJ, Buhr HJ, et al. Rectal anatomy after rectopexy: cinedefecography versus MR-defecography. *Journal of Surgical Research* 2011;**165**(1):52-8.

Pannu 2009 *{published data only (unpublished sought but not used)}*

Pannu HK, Scatarige JC, Eng J. Comparison of supine magnetic resonance imaging with and without rectal contrast to fluoroscopic systocolpoproctography for the diagnosis of pelvic organ prolapse. *Journal of Computer Assisted Tomography* 2009;**33**(1):125-30.

Pescatori 2006 *{published data only (unpublished sought but not used)}*

* Pescatori M, Spyrou M, Pulvirenti d'Urso A. A prospective evaluation of occult disorders in obstructed defecation using the 'iceberg diagram'. *Colorectal Disease* 2006;**8**(9):785-9.

Pescatori M, Spyrou M, Pulvirenti d'Urso A. A prospective evaluation of occult disorders in obstructed defecation using the 'iceberg diagram'. *Colorectal Disease* 2007;**9**(5):452-6.

Pescatori 2009a *{published data only (unpublished sought but not used)}*

Pescatori M, Zbar AP. Reinterventions after complicated or failed STARR procedure. *International Journal of Colorectal Disease* 2009;**24**(1):87-95.

Pescatori 2009b *{published data only (unpublished sought but not used)}*

Pescatori M. Long-term follow-up of simultaneous abdominoperineal repair of enterorectocele and internal mucosal prolapse. *Diseases of the Colon and Rectum* 2009;**52**(2):327-35.

Petersen 2006 *{published data only (unpublished sought but not used)}*

Petersen S, Hellmich G, Schuster A, Lehmann D, Albert W, Ludwig K. Stapled transanal rectal resection under laparoscopic surveillance for rectocele and concomitant enterocele. *Diseases of the Colon and Rectum* 2006;**49**(5):685-9.

Renzi 2016 *{published data only (unpublished sought but not used)}*

Renzi A, Brillantino A, Di Sarno G, d'Aniello F, Bianco P, Iacobellis F, et al. Transverse perineal support: a novel surgical treatment for perineal descent in patients with obstructed defecation syndrome. *Diseases of the Colon and Rectum* 2016;**59**(6):557-64.

Ricchiuti 2016 *{published data only}*

Ricchiuti A, Gambaccini D, Bolognesi V, Costa F, Mumolo M, Ricco G, et al. Dynamic transperineal ultrasound for the evaluation of pelvic floor disorders: which is the best position? *Digestive and Liver Disease* 2016;**48**:e193-e4.

Rizal 2014 *{published data only (unpublished sought but not used)}*

Rizal FE, Suliman I, Vessal S, Shah R, Ahsan S, Mathur P, et al. Investigating symptomatic evacuatory dysfunction: a comparative study between MRI and fluoroscopic defecatory proctograms. *Colorectal Disease* 2014;**16**(Suppl 2):206.

Ron 2018 *{published data only}*

Ron Y. A randomized, open, placebo controlled feasibility study to assess the value of specially designed toilet seat for patients suffering from obstructed defecation type of constipation. *Neurogastroenterology and Motility* 2018;**30 (Supplement 1)**:161.

Schoenenberger 1998 *{published data only (unpublished sought but not used)}*

Schoenenberger AW, Debatin JF, Guldenschuh I, Hany TF, Steiner P, Krestin GP. Dynamic MR defecography with a superconducting, open-configuration MR system. *Radiology* 1998;**206**(3):641-6.

Song 2009 *{published data only (unpublished sought but not used)}*

Song W, Wang Z, Zheng Y, Yi B, Yang X, Jiang T. Application of pelvic floor dynamic MRI combining defecography with homemade high conformable sacculus in the management of obstructed defecation syndrome. *Chung-Hua Wai Ko Tsa Chih [Chinese Journal of Surgery]* 2009;**47**(24):1843-5.

Tsar'kov 2012 *{published data only (unpublished sought but not used)}*

* Tsar'kov PV, Sandrikov VA, Tulina IA, Darinov AA, Brindar NG, Kartaashova OV. Surgical treatment of rectocele with the use of mesh implants by the obstructive defecation syndrome. *Khirurgiia (Mosk)* 2012;**2012**(8):25-33.

Tsarkov PV, Belyaeva l, Lozhkevich A. Anal incontinence and obstructed defecation.Use of dynamic transperineal ultrasound in rectocele and enterocele diagnosis. *Techniques in Coloproctology* 2011;**15**:224.

Wang 2005 *{published data only (unpublished sought but not used)}*

Wang Y, Gong S, Zhang W, Liu B, Zhang L. Comparative study between dynamic MRI and pelvic organography in diagnosis of pelvic floor disorders. *Zhonghua Weichang Waike Zazhi [Chinese Journal of Gastrointestinal Surgery]* 2005;**8**(3):206-9.

Xiong 2006 *{published data only}*

Xiong KL, Gong SG, Zhang WG. Image analysis of puborectalis syndrome and its clinical significance. *Zhonghua Wei Chang Wai Ke za Zhi [Chinese Journal of Gastrointestinal Surgery]* 2006;**9**(6):498-501.

Zeng 2019 *{published data only}*

Guang-Zheng Z, Ben-Qiang R, Yu-Meng L, Shi-Lian Le, Qi-Xu Z, Zhi-Hua P, et al. Application of dynamic MRIdefecography in diagnosis of outletobstructive constipation. *World Chinese Journal of Digestology* 2019;**27**(2):131-8.

Additional references

Altomare 2008

Altomare DF, Spazzafumo L, Rinaldi M, Dodi G, Ghiselli R, Piloni V. Set-up and statistical validation of a new scoring system for obstructed defaecation syndrome. *Colorectal Disease* 2008 Jan;**10**(1):84-8.

Altomare 2018

Altomare DF, Picciariello A, Memeo R, Fanelli M, Digennaro R, Chetta N, et al. Pelvic floor function following ventral rectopexy versus STARR in the treatment of obstructed defecation. *Techniques in Coloproctology* 2018;**22**(4):289-94.

Bartolo 1983

Bartolo DC, Read NW, Jarratt JA, Read MG, Donnelly TC, Johnson AG. Differences in anal sphincter function and clinical presentation in patients with pelvic floor descent. *Gastroenterology* 1983;**85**(1):68-75.

Beggs 2014

Beggs AD, Sultan AH, Thakar R, Abulafi AM. The pelvic floor practice of colorectal surgeons and recommendations for future services. *Bulletin of The Royal College of Surgeons of England* 2014 April;**96**:e1-e8.

Berman 2005

Berman L, Aversa J, Abir F, Longo W. Management of disorders of the posterior pelvic floor. *Yale Journal of Biology and Medicine* 2005;**78**(4):209-18.

Bertschinger 2002

Bertschinger KM, Hetzer FH, Roos JE, Treiber K, Marincek B, Hilfiker PR. Dynamic MR imaging of the pelvic floor performed with patient sitting in an open-magnet unit versus with patient supine in a closed-magnet unit. *Radiology* May 2002;**223**(2):501-8.

Bharucha 2014

Bharucha AE, Rao SSC. An update on anorectal disorders for gastroenterologists. *Gastroenterology* 2014;**146**(1):37-45.

Bock 2013

Bock S, Wolff K, Marti L, Schmied BM, Hetzer FH. Long-term outcome after transanal rectal resection in patients with obstructed defecation syndrome. *Diseases of the Colon and Rectum* 2013;**56**(2):246-52.

Bove 2012

Bove A, Bellini M, Battaglia E, Bocchini R, Gambaccini D, Bove V, et al. Consensus statement AIGO/SICCR diagnosis and treatment of chronic constipation and obstructed defecation (part II: treatment). *World Journal of Gastroenterology* 2012;**18**(36):4994-5013.

Bump 1996

Bump RC, Mattiasson A, Bø K, Brubaker LP, DeLancey JO, Klarskov P, et al. The standardization of terminology of female pelvic organ prolapse and pelvic floor dysfunction. *American Journal of Obstetrics and Gynecology* 1996;**175**(1):10-7.

Chou 2000

Chou Q, Weber AM, Piedmonte MR. Clinical presentation of enterocele. *Obstetrics and Gynecology* Okt 2000;**96**(4):599-603.

Chu 2009

Chu H, Chen S, Louis TA. Random effects models in a meta-analysis of the accuracy of two diagnostic tests without a gold standard. *Journal of the American Statistical Association* 2009;**104**(486):512-23.

Colaiacomo 2009

Colaiacomo MC, Masselli G, Polettini E, Lanciotti S, Casciani E, Bertini L, et al. Dynamic MR imaging of the pelvic floor: a pictorial review. *Radiographics* 2009 May;**29**(3):e35.

Collinson 2008

Collinson R, Cunningham C, D'Costa H, Lindsey I. Rectal intussusception and unexplained faecal incontinence: findings of a proctographic study. *Colorectal Disease* 2009;**11**(1):77-83.

Cronje 2004

Cronje HS, De Beer JA, Bam RH. The pathophysiology of an enterocele and its management. *Journal of Obstetrics and Gynaecology* 2004;**24**(4):408-13.

D'Hoore 2003

D?Hoore A, Penninckx F. Obstructed defecation. *Colorectal Disease* 2003;**5**(4):280-7.

Deeks 2005

Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. *Journal of Clinical Epidemiology* 2005;**58**(9):882-93.

DeLancey 1999

DeLancey JO. Structural anatomy of the posterior pelvic compartment as it relates to rectocele. *American Journal of Obstetrics and Gynecology* 1999;**180**(4):815-23.

Dennison 2005

Dennison C, Prasad M, Lloyd A, Bhattacharyya SK, Dhawan R, Coyne K. The health-related quality of life and economic burden of constipation. *PharmacoEconomics* 2005;**23**(5):461-76.

Dietz 2005a

Dietz HP, Steensma AB. Posterior compartment prolapse on two-dimensional and three-dimensional pelvic floor ultrasound: the distinction between true rectocele, perineal hypermobility and enterocele. *Ultrasound in Obstetrics & Gynecology* 2005;**26**(1):73-7.

Dietz 2005b

Dietz HP, Clarke B. Prevalence of rectocele in young nulliparous women. *Australian & New Zealand Journal of Obstetrics & Gynaecology* 2005;**45**(5):391?4.

Dietz 2006

Dietz HP, Steensma AB. The role of childbirth in the aetiology of rectocele. *British Journal of Obstetrics and Gynaecology* 2006;**113**(3):264?7.

Dietz 2012

Dietz HP, Beer-Gabel M. Ultrasound in the investigation of posterior compartment vaginal prolapse and obstructed defecation. *Ultrasound in Obstetrics & Gynecology* 2012;**40**(1):14-27.

Dietz 2014

Dietz HP. Translabial ultrasound in the assessment of pelvic floor and anorectal function in women with defecatory disorders. *Techniques in Coloproctology* 2014;**18**(5):481-94.

Ferrante 1991

Ferrante SL, Perry RE, Schreiman JS, Cheng SC, Frick MP. The reproducibility of measuring the anorectal angle in defecography. *Diseases of the Colon and Rectum* 1991;**34**(1):51-5.

Fielding 1998

Fielding JR, Griffiths DJ, Versi E, Mulkern RV, Lee ML, Jolesz FA. MR imaging of pelvic floor continence mechanisms in the supine and sitting positions. *American Journal of Roentgenology* 1998;**171**(6):1607-10.

Freimanis 1991

Freimanis MG, Wald A, Caruana B, Bauman DH. Evacuation proctography in normal volunteers. *Investigative Radiology* 1991;**26**(6):581-5.

Garrigues 2004

Garrigues V, Gálvez C, Ortiz V, Ponce M, Nos P, Ponce J. Prevalence of constipation: agreement among several criteria and evaluation of the diagnostic accuracy of qualifying symptoms and self-reported definition in a population-based survey in Spain. *American Journal of Epidemiology* 2004;**159**(5):520-6.

Goei 1989

Goei R, Van Engelshoven J, Schouten H, Baeten C, Stassen C. Anorectal function: defecographic measurement in asymptomatic subjects. *Radiology* 1989;**173**(1):137-41.

Grossi 2018

Grossi U, Di Tanna GL, Heinrich H, Taylor SA, Knowles CH, Scott SM. Systematic review with meta-analysis: defecography should be a first-line diagnostic modality in patients with refractory constipation. *Alimentary Pharmacology & Therapeutics* 2018;**48**(11-12):1186-201.

Guyatt 2008

Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al, GRADE Working Group. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. *BMJ* 2008;**336**(7650):924-6.

Guzman Rojas 2016

Guzman Rojas R, Kamisan Atan I, Shek KL, Dietz HP. The prevalence of abnormal posterior compartment anatomy and its association with obstructed defecation symptoms in urogynecological patients. *International Urogynecology Journal* 2016;**27**(6):939-44.

Halligan 1995

Halligan S, Bartram CI, Park HJ, Kamm MA. Proctographic features of anismus. *Radiology* 1995;**197**(3):679-82.

Horne 2019

Horne DJ, Kohli M, Zifodya JS, Schiller I, Dendukuri N, Tollefson D, et al. Xpert MTB/RIF and Xpert MTB/RIF Ultra for pulmonary tuberculosis and rifampicin resistance in adults. *Cochrane Database of Systematic Reviews* 2019, Issue 6.

Hsu 2011

Hsu J, Bro?ek JL, Terracciano L, Kreis J, Compalati E, Stein AT, et al. Application of GRADE: making evidence-based recommendations about diagnostic tests in clinical practice guidelines. *Implementation Science* 2011;**6**:62.

Irvine 2002

Irvine EJ, Ferrazzi S, Pare P, Thompson WG, Rance L. Health-related quality of life in functional GI disorders: focus on constipation and resource utilization. *American Journal of Gastroenterology* 2002;**97**(8):1986-93.

Kahn 2005

Kahn MA, Breitkopf CR, Valley MT, Woodman PJ, O?Boyle AL, Bland DI, et al. Pelvic Organ Support Study (POSST) and bowel symptoms: Straining at stool is associated with perineal and anterior vaginal descent in a general gynecologic population. *American Journal of Obstetrics and Gynecology* 2005;**192**(5):1516?22.

Kapoor 2008

Kapoor DS, Sultan AH, Thakar R, Abulafi MA, Swift RI, Ness W. Management of complex pelvic floor disorders in a multidisciplinary pelvic floor clinic. *Colorectal Disease* 2008;**10**(2):118-23.

Karasick 1997

Karasick S, Spettell CM. The role of parity and hysterectomy on the development of pelvic floor abnormalities revealed by defecography. *American Journal of Roentgenology* 1997;**169**(6):1555-8.

Karlbom 1999

Karlbom U, Nilsson S, Påhlman L, Graf W. Defecographic study of rectal evacuation in constipated patients and control subjects. *Radiology* 1999;**210**(1):103-8.

Kelvin 1992

Kelvin FM, Maglinte DD, Hornback JA, Benson JT. Pelvic prolapse: assessment with evacuation proctography (defecography). *Radiology* 1992;**184**(2):547-51.

Kelvin 1999

Kelvin FM, Hale DS, Maglinte DD, Patten BJ, Benson JT. Female pelvic organ prolapse: diagnostic contribution of dynamic cystoproctography and comparison with physical examination. *American Journal of Roentgenology* 1999;**173**(1):31-7.

Kleeman 2005

Kleeman SD, Westermann C, Karram MM. Rectoceles and the anatomy of the posterior vaginal wall: revisited. *American Journal of Obstetrics and Gynecology* 2005;**193**(6):2050-5.

Kuijpers 1985

Kuijpers HC, Bleijenberg G. The spastic pelvic floor syndrome. A cause of constipation. *Diseases of the Colon and Rectum* 1985;**28**(9):669-72.

Lapalus 2004

Lapalus MG, Henry L, Barth X, Mellier G, Gautier G, Mion F, et al. Enterocele: clinical risk factors and association with others pelvic floor disorders (about 544 defecographies). *Gynecologie, Obstetrique & Fertilite* 2004;**32**(7-8):595?600.

Lembo 2003

Lembo A, Camilleri M. Chronic constipation. *New England Journal of Medicine* 2003;**349**(14):1360-8.

Leroi 1995

Leroi AM, Berkelmans I, Denis P, Hémond M, Devroede G. Anismus as a marker of sexual abuse. Consequences of abuse on anorectal motility. *Digestive Diseases and Sciences* 1995;**40**(7):1411-6.

Ludwikowski 2002

Ludwikowski B, Hayward IO, Fritsch H. Rectovaginal fascia: an important structure in pelvic visceral surgery? About its development, structure, and function. *Journal of Pediatric Surgery* 2002;**37**(4):634-8.

Lunn 2009

Lunn D, Spiegelhalter D, Thomas A, Best N. The BUGS project: Evolution, critique and future directions. *Statistics in Medicine* 2009;**28**(25):3049-67.

Mahieu 1984

Mahieu P, Pringot J, Bodart P. Defecography: I. Description of a new procedure and results in normal patients. *Gastrointestinal Radiology* 1984;**9**(3):247-51.

Mellgren 1994a

Mellgren A, Bremmer S, Johansson C, Dolk A, Udén R, Ahlbäck SO, et al. Defecography. Results of investigations in 2,816 patients. *Diseases of the Colon and Rectum* 1994;**37**(11):1133-41.

Mellgren 1994b

Mellgren A, Johansson C, Dolk A, Anzén B, Bremmer S, Nilsson BY, et al. Enterocele demonstrated by defaecography is associated with other pelvic floor disorders. *International Journal of Colorectal Disease* 1994;**9**(3):121-4.

Menten 2015

Menten J, Lesaffre E. A general framework for comparative Bayesian meta-analysis of diagnostic studies. *BMC Medical Research Methodology* 2015;**15**:70.

Morandi 2010

Morandi C, Martellucci J, Talento P, Carriero A. Role of enterocele in the obstructed defecation syndrome(ODS): a new radiological point of view. *Colorectal Disease* 2010;**12**(8):810-6.

Murad-Regadas 2012a

Murad-Regadas SM, Rodrigues LV, Furtado DC, Regadas FS, Olivia da S Fernandes G, Regardas Filho FS, et al. The influence of age on posterior pelvic floor dysfunction in women with obstructed defecation syndrome. *Techniques in Coloproctology* 2012;**16**(3):227-32.

Müller-Lissner 1998

Müller-Lissner SA, Bartolo DC, Christiansen J, Ekberg O, Goei R, Höpfner W. Interobserver agreement in defecography? an international study. *Zeitschrift fur Gastroenterologie* 1998;**36**(4):273-9.

Nichols 1972

Nichols DH. Types of enterocele and principles underlying choice of operation for repair. *Obstetrics and Gynecology* 1972;**40**(2):257-63.

Nisenblat 2016

Nisenblat V, Prentice L, Bossuyt PMM, Farquhar C, Hull ML, Johnson N. Combination of the non-invasive tests for the diagnosis of endometriosis. *Cochrane Database of Systematic Reviews* 2016, Issue 7. Art. No: CD012281. [DOI: 10.1002/14651858.CD012281]

Noelting 2016

Noelting J, Eaton JE, Choung RS, Zinsmeister AR, Locke GR, Bharucha AE. The incidence rate and characteristics of clinically diagnosed defecatory disorders in the community. *Neurogastroenterology and Motility* 2016;**28**(11):1690-7.

Oom 2009

Oom DJ, Gosselink MP, Schouten WR. Enterocele - diagnosis and treatment. *Gastroenterologie Clinique et Biologique* 2009;**33**(2):135-7.

Palit 2014

Palit S, Bhan C, Lunniss PJ, Boyle DJ, Gladman MA, Knowles CH, et al. Evacuation proctography: a reappraisal of normal variability. *Colorectal Disease* 2014;**16**(7):538-46.

Papatheodoridis 2010

Papatheodoridis GV, Vlachogiannakos J, Karaitianos I, Karamanolis DG. A Greek survey of community prevalence and characteristics of constipation. *European Journal of Gastroenterology & Hepatology* 2010;**22**(3):354-60.

Pare 2001

Pare P, Ferrazzi S, Thompson WG, Irvine EJ, Rance L. An epidemiological survey of constipation in Canada: definitions, rates, demographics, and predictors of health care seeking. *American Journal of Gastroenterology* 2001;**96**(11):3130-7.

Parks 1966

Parks AG, Porter NH, Hardcastle J. The syndrome of the descending perineum. *Proceedings of the Royal Society of Medicine* June 1966;**59**(6):477-82.

Piloni 2013

Piloni V, Tosi P, Vernelli M. MR-defecography in obstructed defecation syndrome (ODS): technique, diagnostic criteria and grading. *Techniques in Coloproctology* 2013;**17**(5):501-10.

Pizzoferrato 2014

Pizzoferrato AC, Nyangoh Timoh K, Fritel X, Zareski E, Bader G, Fauconnier A. Dynamic Magnetic Resonance Imaging and pelvic floor disorders: how and when? *European Journal of Obstetrics, Gynecology, and Reproductive Biology* 2014;**181**:259-66.

Plummer 2006

Plummer M, Best N, Cowles K, Vines K. CODA: Convergence Diagnosis and Output Analysis for MCMC. *R News* 2006;**6**:7-11.

Podzemny 2015

Podzemny V, Pescatori LC, Pescatori M. Management of obstructed defecation. *World Journal of Gastroenterology* 2015;**21**(4):1053-60.

R Core Team 2019 [Computer program]

R Foundation for Statistical Computing R: A language and environment for statistical computing. R Core Team. Vienna, Austria: R Foundation for Statistical Computing, 2019. www.R-project.org/.

Racaniello 2015

Racaniello E, Terzoni S, Accardi R, Ricci C, Boccasanta P, Destrebecq A. Quality of life of patients undergoing surgery for obstructed defecation syndrome: A before-after study. *International Journal of Surgery* 2015;**21**:18-21.

Ramage 2017

Ramage L, Simillis C, Yen C, Lutterodt C, Qiu S, Tan E, et al. Magnetic resonance defecography versus clinical examination and fluoroscopy: a systematic review and meta-analysis. *Techniques in Coloproctology* 2017;**21**(12):915-27.

Reitsma 2005

Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, Zwinderman AH. Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. *Journal of Clinical Epidemiology* 2005;**58**:982-90.

Renzi 2013

Renzi A, Brillantino A, Di Sarno G, D'Aniello F. Five-Item score for obstructed defecation syndrome: study of validation. *Surgical Innovation* 2013;**20**(2):119-25.

Richardson 1993

Richardson AC. The rectovaginal septum revisited: its relationship to rectocele and its importance in rectocele repair. *Clinical Obstetrics and Gynecology* 1993;**36**(4):976-83.

Rodrigo 2011

Rodrigo N, Shek KL, Dietz HP. Rectal intussusception is associated with abnormal levator ani muscle structure and morphometry. *Techniques in Coloproctology* 2011;**15**(1):39-43.

Roos 2002

Roos J, Weishaupt D, Wildermuth S, Willmann JK, Marincek B, Hilfiker PR. Experience of 4 years with open MR defecography: pictorial review of anorectal anatomy and disease. *Radiographics* 2002;**22**(4):817-32.

Rutjes 2007

Rutjes AW, Reitsma JB, Coomarasamy A, Khan KS, Bossuyt PM. Evaluation of diagnostic tests when there is no gold standard. A review of methods. *Health Technology Assessment* 2007;**11**(iii):ix?51.

Santoro 2006

Santoro GA, Di Falco G. Update in the evaluation of outlet obstruction. In: Benign Anorectal Diseases; Diagnosis with Endoanal and Endorectal Ultrasound and New Treatment Options. Springer Verlag, 2006:205-58.

Santoro 2011

Santoro GA, Wieczorek AP, Dietz HP, Mellgren A, Sultan AH, Shobeiri SA, et al. State of the art: an integrated approach to pelvic floor ultrasonography. *Ultrasound in Obstetrics & Gynecology* 2011;**37**(4):381-96.

Shobeiri 2012

Shobeiri SA, White D, Quiroz LH, Nihira MA. Anterior and posterior compartment 3D endovaginal ultrasound anatomy based on direct histologic comparison. *International Urogynecology Journal* 2012;**23**(8):1047-53.

Shorvon 1989

Shorvon PJ, McHugh S, Diamant NE, Somers S, Stevenson GW. Defecography in normal volunteers: results and implications. *Gut* 1989;**30**(12):1737-49.

Singh 2012

Singh S, Chang SM, Matchar DB, Bass EB. Chapter 7: Grading a body of evidence on diagnostic tests. In: Methods Guide for Medical Test Reviews. Agency for Healthcare Research and Quality, 2012.

Stoker 2000

Stoker J, Rociu E, Wiersma TG, Laméris JS. Imaging of anorectal disease. *British Journal of Surgery* 2000;**87**(1):10-27.

Stoker 2001

Stoker J, Halligan S, Bartram CI, . Pelvic floor imaging. *Radiology* 2001;**218**(3):621-41.

Sturtz 2005

Sturtz S, Ligges U, Gelman A. R2WinBUGS: A Package for Running WinBUGS from R. *Journal of Statistical Software* 2005;**12**(3):1-16.

Sultan 2017

Sultan AH, Monga A, Lee J, Emmanuel A, Norton C, Santoro G, et al. An International Urogynecological Association (IUGA)/ International Continence Society (ICS) joint report on the terminology for female anorectal dysfunction. *International Urogynecology Journal* 2017;**28**(1):5-31.

Takahashi 2006

Takahashi T, Yamana T, Sahara R, Iwadare J. Enterocele: what is the clinical implication? *Diseases of the Colon and Rectum* 2006;**49**(10 (Suppl)):S75-81.

Tan 2005

Tan JS, Lukacz ES, Menefee SA, Powell CR, Nager CW. Predictive value of prolapse symptoms: a large database study. *International Urogynecology Journal* 2005;**16**(3):203?9.

Thomas 2006

Thomas A, O'Hara B, Ligges U, Sturtz S. Making BUGS Open. *R News* 2006;**6**(1):12-17.

Whitcomb 2009

Whitcomb EL, Lukacz ES, Lawrence JM, Nager CW, Luber KM. Prevalence of defecatory dysfunction in women with and without pelvic floor disorders. *Journal of Pelvic Medicine and Surgery* 2009;**15**(4):179-217.

Whiting 2011a

Whiting P, Westwood M, Beynon R, Burke M, Sterne JA, Glanville J. Inclusion of methodological filters in searches for diagnostic test accuracy studies misses relevant studies. *Journal of Clinical Epidemiology* 2011;**64**(6):602-7.

Whiting 2011b

Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of Internal Medicine* 2011;**155**(8):529-36.

Whiting 2013

Whiting PF, Rutjes AW, Westwood ME, Mallett S. A systematic review classifies sources of bias and variation in diagnostic test accuracy studies. *Journal of Clinical Epidemiology* 2013;**66**(10):1093-104.

Wieczorek 2011

Wieczorek AP, Stankiewicz A, Santoro GA, Wo?niak MM, Bogusiewicz M, Rechberger T. Pelvic floor disorders: role of new ultrasonographic techniques. *World Journal of Urology* 2011;**29**(5):615-23.

Williams 2010

Williams PM, Fletcher S. Health effects of prenatal radiation exposure. *American Family Physician* 2010;**82**(5):488-93.

Yoshioka 1991

Yoshioka K, Matsui Y, Yamada O, Sakaguchi M, Takada H, Hioki K, et al. Physiologic and anatomic assessment of patients with rectocele. *Diseases of the Colon and Rectum* 1991;**34**(8):704-8.

Zbar 2003

Zbar AP, Lienemann A, Fritsch H, Beer-Gabel M, Pescatori M. Rectocele: pathogenesis and surgical management. *International Journal of Colorectal Disease* 2003;**18**(5):369?84.

* Indicates the major publication for the study

CHARACTERISTICS OF STUDIES

Characteristics of included studies *[ordered by study ID]*

Barthet 2000

Study characteristics	
Patient Sampling	**Patient selection:** A total of 43 participants were prospectively enrolled in the present study, between February 1997 and September 1998. All had symptoms involving outlet delay with either exaggerated effort during defaecation, manual disimpaction of stool, vaginal manoeuvres to aid defaecation, or a feeling of incomplete defaecation **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To determine the accuracy of dynamic anorectal endosonography (DAE) as compared with defaecography as a means of assessing pelvic floor disorders **Inclusion criteria:** Women with symptoms involving outlet delay **Exclusion criteria:** Not described
Patient characteristics and setting	**Nr of included patients:** 43 **Gender:** 43 women (100%) **Age:** mean age 51, range 30 - 74

Barthet 2000 *(Continued)*

	Symptoms: All had symptoms involving outlet delay **Ethnicity:** Not described **Co-morbidities:** 32 had undergone at least 1 vaginal delivery, with 13 cases of perianal tear, 10 cases of episiotomy, and 4 cases of forceps delivery. A total of 12 participants had previously undergone hysterectomy (10 abdominal, 2 vaginal), and 19 complained of urinary stress incontinence **Setting:** Secondary care, single centre **Time period:** Between February 1997 and September 1998 **Country study is conducted:** France
Index tests	**Name index test:** Dynamic anal endosonography (DAE) **Details of conducting index test:** DAE was performed on participants lying in the left lateral decubitus position. A rigid linear endoanal probe with a frequency of 7MHz was used (model PVL-625RT; Toshiba, Tokyo, Japan). The probe was 20 mm wide and 120 mm long, with a field View of 57 mm. The tip of the probe was covered with a water-filled balloon to maintain the acoustic coupling **Imaging acquisition:** By rotating the linear probe through 360 °, the various layers constituting the anal wall (mucosa, internal sphincter, external sphincter), the layer forming the rectal wall, and the perirectal tissues (puborectal muscle, bladder, vagina, or prostate) could be identified. At the end of the initial examination, the participant was asked to produce a defaecation effort while anal ultrasonography was continued, leaving the ultrasound probe in the same position. In the last 15 participants the rectum was filled with 50 ml water before defaecation effort **Imaging analysis:** The descent of the puborectal muscle during defaecation effort were measured, the same procedure being repeated 3 times. The position of the puborectal muscle was first marked at rest with a fine calliper. The puborectal muscle appeared in the form of an oblique hypoechoic layer, with a fine hyperechoic line connecting it to the anal canal at the level of the anorectal angle. The participant was then asked to produce a straining effort, and the new position of the puborectal muscle was recorded. The descent of the puborectal muscle corresponded to the distance between its initial position and its position at the end of the straining effort. Rectocele was identified by the mobilisation of air during the straining effort; this procedure was improved by filling of the rectum with water. All the parameters were analysed by the same operator. All the investigations were carried out without prior knowledge of previous findings **Threshold test positivity:** Rectocele present/absent, Intussusception present/absent, pelvic floor descent present/absent
Target condition and reference standard(s)	**Name of index test 'EP':** Defaecography **Details of conducting** evacuation proctography**:** Defaecography was performed using a simplified method described by Mahieu et al. After sufficient contrast filling of the rectum, the participant was asked to sit on a special commode. No opacification of the small bowel or the vagina was performed **Imaging acquisition:** The participant was asked to contract the pelvic floor musculature and then to empty the rectum as completely as possible. The fluoroscopic images were recorded during several such manoeuvres, in order to assess and measure the descent of the pelvic floor and to diagnose any rectocele or rectal intussusception **Imaging analysis:** All the parameters were analyses by the same radiologist. All the assessments were recorded under blinded conditions on separate sheets **Threshold test positivity:** Rectocele: present/absent, Intussusception: full thickness circumferential infolding present/absent, perineal descent: > 2 cm below the pubococcygeal line at rest or descent to > 3 cm below the pubococcygeal line on straining
Flow and timing	**Enrolment and exclusions (+ reasons):** DAE and defaecography were performed on all the participants **Nr analysed:** 43

Barthet 2000 *(Continued)*

Time interval (+ interventions) between index test and reference standard: Not described

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Unclear		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Barthet 2000 *(Continued)*

DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Beer-Gabel 2004

Study characteristics	
Patient Sampling	**Patient selection:** 33 consecutive women who presented with long-standing difficulty in evacuation were examined with both proctography and DTP-US **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The purpose of this study was to assess the level of agreement between evacuation proctography and DTPUS in diagnosing pathology in an unselected group of participants who presented to our pelvic floor clinic with evacuatory difficulty and to compare measurements of anorectal configuration using both the techniques **Inclusion criteria:** Women with longstanding reported history of constipation (history of evacuatory difficulty exceeded 6 months)

Beer-Gabel 2004 *(Continued)*

	Exclusion criteria: *Additional information from authors:* Patients who did not have both examinations
Patient characteristics and setting	**Nr of included patients:** 33 **Gender:** Female (100%) **Age:** Mean age 58 years, range 32 – 77 **Symptoms:** Only patients defined as constipated were included in the study, in which their history of evacuatory difficulty exceeded 6 months, if there was at most 1 bowel movement every 4 days (or longer), and/or if more than 25% of these movements were accompanied by excessive straining. Of the selected participants 14 (42.4%) complained of daily straining at stool, with 8 (24.2%) reporting only 1 stool on average per week. 15 participants (46%) complained of repeatedly unsatisfied defaecation, with 7 (21.2%) reporting hard stools for more than 50% of evacuation attempts. 16 participants (48.5%) used daily stimulant laxatives, with 4 (12.1%) using daily enemas and a further 4 admitting to daily rectal digitation to assist evacuation. 8 participants (24.2%) confirmed that toileting exceeded 60 minutes duration on average **Ethnicity:** *Additional information from authors:* white **Co-morbidities:** 12 participants in the group (36.4%) had previously undergone a hysterectomy, 6 (18.2%) a haemorrhoidectomy, and 3 (9.1%) lateral internal anal sphincterotomy, with a further 3 (9.1%) having a confirmed solitary rectal ulcer **Setting:** Tertiary care, single centre **Time period:** 2003 **Country study is conducted:** Israel
Index tests	**Name index test:** Dynamic Transperineal Ultrasound (DTP-US) **Details of conducting index test:** All procedures were videotaped for orthograde and retrograde scrolling of dynamic images and static representative images were used for clinical measurement. DTP-US was performed using curvilinear transducers (C 4-7 and C 8-12) and a linear-array transducer (L 5-10 ATL, HDI 3000, Advanced Technology Laboratories, Bothell, Wash., USA). Before the start of the procedure the participant's rectum was filled with 50 ml ultrasonographic coupling gel (Ultra-Gel, Aquarius 101, Medilab USA) using a standard Luer syringe and a soft-end catheter. A similar volume of acoustic gel was instilled into the vagina and gel was liberally applied to the perineum. Participants were advised to avoid micturition for a minimum of 2 hours prior to the procedure. Gastrografin (50 ml) diluted 1:1 with tap water was ingested by the participant 1 hour prior to each procedure. The perineum of the participant was examined in the left-lateral position **Imaging acquisition/analysis:** Images of the infra levator viscera and soft tissues and the pelvic floor musculature were obtained at rest and during maximal straining for routine visualisation of the pubis, urethra, bladder, vagina, anus, distal rectum, and puborectalis muscle, all of which were registered by the examiner. All examinations were performed by the same clinician (M.B.G.) who was blinded to the results of defaecography. Sagittal examination of the anterior perineum showed the distal vagina, bladder, and urethra and was used to identify contrast-filled enteric loops (if present) between the rectal and vaginal walls in the territory of the rectovaginal septum. Towards the end of the procedure participants were encouraged to evacuate as much of the intrarectal gel as possible **Threshold test positivity:** Unknown
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography:** Evacuation proctography was performed without prior bowel preparation with 120 ml barium paste (55% wt/wt barium sulfate) instilled into the rectum using conventional video-fluoroscopy in the lateral sitting position at rest and during evacuation in accordance with the basic technique described by Shorvon 1989. The small bowel was opacified following ingestion of 200 ml dilute oral diatrizoate meglumine (Gastrografin, Schering UK) 60 minutes prior to the examination **Imaging acquisition/analysis:** All proctographic examinations and measurements were made by the same examiner (A.P.Z.) who was blinded to the results of DTP-US

Beer-Gabel 2004 *(Continued)*

	Threshold test positivity: Unknown
Flow and timing	**Enrolment and exclusions (+ reasons):** All included women received both investigations and were included in the 2x2 table. **Nr analysed:** 33 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* Less than 2 weeks
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Beer-Gabel 2004 *(Continued)*

If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Beer-Gabel 2008

Study characteristics	
Patient Sampling	**Patient selection:** 62 consecutive women referred to a specialised Pelvic Floor Unit, Chaim Sheba Tel-Hashomer Hospital, Israel with long-standing symptoms of obstructed defaecation were assessed for analysis in this study **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** This study compares DTP-US with DEP specifically for the diagnosis of cul-de-sac hernias among patients presenting to a specialised pelvic-floor-dysfunction clinic principally with obstructed defaecation **Inclusion criteria:** Long-standing symptoms of obstructed defaecation **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 62 **Gender:** Female 100% **Age:** Mean: 56.2; Range: 21 – 90 **Symptoms:** The clinical diagnosis of obstructed defaecation was considered when the participants needed to strain in evacuation more than 25% of the time, in accordance with the ROME II criteria, and when there was an attendant feeling of incomplete defaecation, repetitive attempts to defaecate, and where these symptoms exceeded 6 months duration. In 24 participants (38.7%), obstructed defaecation was part of an irritable bowel syndrome, whereas 22 participants (35.5%) had functional constipation, and 16 (25.8%) concomitant faecal incontinence associated with their constipation **Ethnicity:** *Additional information from authors:* white **Co-morbidities:** 18 participants (30%) had undergone a prior hysterectomy with 13 (22%) having previous abdominal surgery and 5 (8%), anal surgery **Setting:** Tertiary care, single centre **Time period:** Between August 2004 and October 2005 **Country study is conducted:** Israel
Index tests	**Name index test:** Dynamic transperineal ultrasonography (DTP-US) **Details of conducting index test:** All examinations with DTP-US were performed by one of the authors (MBG) blinded to the DEP results. DTP-US was performed in accordance with our prior reported technique using either a curvilinear C4–7 or a C8–12 transducer (Logiq 9, GE Healthcare UK). The transducer was protected with a latex condom, and images were routinely obtained from structures in the anterior compartment, (the pubis, urethra, and bladder), the middle compartment, (the vagina and the rectovaginal septum) and the posterior compartment, (the anal canal, the rectum, and the puborectalis muscle en face). Before the performance of the DTP-US, the rectum was instilled with 50 mL of ultrasonographic coupling gel (Ultragel Aquarius 101® Medilab, USA) using a standard Luer syringe with a soft-end catheter. Opacification of the vagina was routinely performed with 20 mL of acoustic gel. The participants were advised to avoid micturition for a 1-hour period before the procedure, and 50 mL of Gastrografin (diluted 1:1 with tap water) was ingested by the participant 1 h before the DTP-US. **Imaging acquisition:** The images were obtained in the mid-sagittal plane and at various transverse points of the posterior compartment and the perineal body at the mid-anal canal level at rest, during maximal straining and squeeze, and in some cases, during rectal evacuation **Imaging analysis:** Unknown **Threshold test positivity:** Enteroceles were readily identified as small bowel loops visible in the region of the rectovaginal septum. Peritoneoceles were defined as an enlarged rectovaginal septum without visible small-bowel loops being present.

Beer-Gabel 2008 *(Continued)*

Target condition and reference standard(s)	**Name of index test 'EP':** Dynamic evacuation proctography (DEP) **Details of conducting evacuation proctography:** DEP was performed by 2 investigators (AY and MA) blinded to the clinical and DTP-US results. Participants were given 10 mL of Gastrografin (Schering ®, UK) diluted with 150 mL of tap water and 50 mL of barium 30 minutes before the performance of the DEP to opacify the small bowel. The distal colon and rectum were filled with 150 mL of contrast medium using a mixture of barium with oatmeal powder (140 mL of barium sulphate with 20 g of oatmeal) so as to obtain a stool-like consistency. The vagina was opacified with 20 mL of barium paste **Imaging acquisition:** The participant was then seated on a dedicated commode with films being obtained at rest, during squeeze, and during maximal straining in accordance with standard techniques. Both static views and video records were made for each participant **Imaging analysis:** Unknown **Threshold test positivity:** Enteroceles were diagnosed when a loop or loops of small bowel were detected in the territory between the rectum and the vagina, compressing the anterior rectal wall
Flow and timing	**Enrolment and exclusions (+ reasons):** All patients received index test and reference standard. **Nr analysed:** 62 **Time interval (+ interventions) between index test and reference standard:** Unclear
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		

Beer-Gabel 2008 *(Continued)*

Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Beer-Gabel 2008 *(Continued)*

Could the patient flow have introduced bias?	Unclear risk

Beer-Gabel 2015

Study characteristics	
Patient Sampling	**Patient selection:** This study is a retrospective review of data that had been collected in a population of women that were referred to our clinic for the evaluation of evacuation disorders (chronic constipation and faecal incontinence) during the years 2011 to 2013. Symptom severity for faecal incontinence was determined based on the Wexner score. In order to include only women with significant symptoms, we included in the faecal incontinence group only participants who had solid or liquid faecal incontinence more than once a month. We assessed chronic constipation symptoms based on the Cleveland Constipation Severity Index (SCCI) scoring system. We included in the constipation group only participants who scored 15 or higher. As part of their evaluation, all women were examined by dynamic transperineal ultrasonography (DTP-US) and defaecography (DEF) **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** The aim of this study was to evaluate the level of consistency between DEF and DTP-US in the diagnosis of pelvic floor deformations **Inclusion criteria:** Women with chronic constipation or faecal incontinence **Exclusion criteria:** Wexner score: solid or liquid faecal incontinence less than once a month. Cleveland Constipation Severity Index (SCCI) scoring system: scored lower than 15
Patient characteristics and setting	**Nr of included patients:** 105 **Gender:** Female 100% **Age:** 54.6 ± 11 years **Symptoms:** 81 women were evaluated for chronic constipation and 24 for faecal incontinence **Ethnicity:** *Additional information from authors:* White **Co-morbidities:** Mean parity was 2.8 ± 1.2 **Setting:** Tertiary care, single centre **Time period:** 2011 - 2013 **Country study is conducted:** Israel
Index tests	**Name index test:** Dynamic transperineal ultrasound (DTP-US) **Details of conducting index test:** DTP-US is readily performed after rectal cleansing with 1 enema. DTP-US was conducted using a curvilinear 5–8 MHz (B&K, Profocus Ultra View, Herlev, Denmark) probe after liberal application of acoustic gel to the perineum and instilling 10 ml gel intravaginally and 50 ml gel into the rectum **Imaging acquisition:** The examination was then performed while the participant was lying in the left lateral position at rest, at squeeze, and at straining **Imaging analysis:** All exams were performed and interpreted by a single physician (MBG) who was blinded to the results of the DEF

Beer-Gabel 2015 *(Continued)*

	Threshold test positivity: Rectocele: > 2 cm; enterocele: any; intussusception: any; anismus: any; perineal descent: > 2 cm
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography & imaging acquisition:** Evacuation proctography was performed without prior bowel preparation with 120 ml barium paste (55 % wt/wt barium sulfate) instilled into the rectum using conventional videofluoroscopy while the participant was in the lateral sitting position at rest and during evacuation in accordance with the basic technique described previously. The small bowel was opacified following ingestion of 100 ml dilute oral diatrizoate meglumine (Gastrografin, Schering UK) 40 minutes prior to the examination **Imaging analysis:** All proctographic examinations and measurements were done by a single examiner who was blind to the results of the DTP-US **Threshold test positivity:** Rectocele: depth > 2 cm; enterocele: any; intussusception: any; anismus: any; perineal descent: descent ARJ > 2 cm
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants were examined by DTP-US and DEF and included in the 2x2 table **Nr analysed:** 105 **Time interval (+ interventions) between index test and reference standard:** Unclear
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge	Yes		

Beer-Gabel 2015 *(Continued)*

of the results of the other index test(s)?			
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Brusciano 2007

Study characteristics	
Patient Sampling	**Patient selection:** From a prospectively-collected database, 92 consecutive patients with symptoms of OD (straining at stool, sense of incomplete evacuation, need for self-digitations, use of laxatives) seen in our units over a 3-year period were evaluated **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** The purpose of our study was to investigate the findings of anal–vaginal–dynamic perineal US in patients with obstructed defaecation compared to healthy controls, to correlate them with manometry and defaecography and to evaluate their clinical usefulness in participants who are candidates for surgery **Inclusion criteria:** Patients with symptoms of ODS **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 92 **Gender:** Women 77 (84%), men 15 (16%) **Age:** Mean age 51.3 ± 11 years; range 21 – 71 **Symptoms:** Symptoms of OD (straining at stool, sense of incomplete evacuation, need for self-digitations, use of laxatives) **Ethnicity:** *Additional information from authors:* white **Co-morbidities:** Unknown **Setting:** Secondary, single centre **Time period:** 2003 - 2006 **Country study is conducted:** Italy
Index tests	**Name index test:** Dynamic perineal ultrasound **Details of conducting index test:** Dynamic perineal US was performed in the gynaecologic position using a linear 5- to 8-mHz probe. Transverse images were obtained by placing the probe on the perineum, between the anus and the introit. The longitudinal section of the anal canal and the puborectalis sling may be scanned by changing application pressure and probe inclination **Imaging acquisition:** Images were acquired at rest and during straining **Imaging analysis:** Unknown **Threshold test positivity:** *Additional information from the authors:* Rectocele: > 1 cm depth; enterocele: bowel loops in rectovaginal space; intussusception: any; anismus: more acute angle on straining
Target condition and reference standard(s)	**Name index test 'EP':** Defaecography **Details of conducting evacuation proctography and imaging acquisition:** Defaecography was performed by introducing a barium paste in the rectum and taking radiographs at resting, squeezing and straining as reported by Mahieu 1984. Enterocolpodefaecography aimed at detecting an enterocele/sigmoidocele was performed in 28 of the 43 participants **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: > 2 cm; enterocele: small bowels in rectovaginal space; intussusception: any; anismus: lack of shortening and widening of the anal canal on straining

Brusciano 2007 *(Continued)*

Flow and timing	**Enrolment and exclusions (+ reasons):** Defaecography was performed in 43 (47%) of 92 participants, of which entero-colpo-defaecography was performed in 28 aimed at detecting an enterocele/sigmoidocele. Dynamic perineal US was carried out in 41 (44%) of the 92 participants **Nr analysed:** 43 women had defaecography and 41 defaecography and dynamic perineal ultrasound **Time interval (+ interventions) between index test and reference standard:** Unknown
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Brusciano 2007 *(Continued)*

DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	No		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Dellemare 1994

Study characteristics	
Patient Sampling	**Patient selection:** 33 consecutive women with suspected anterior rectocele (ARC) were subjected to radiographic defaecography. Selection criteria for ARC-related complaints in these women were the need to support the anterior rectal wall digitally during evacuation, incomplete evacuation, false urgency, a feeling of outlet obstruction, faecal incontinence (also if this occurs during coitus), and a feeling of perineal fatigue, perineal pressure, or vaginal prolapse **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of this study was to devise a measuring method for an anterior rectocele on standardised defaecographies and magnetic resonance images (MRI) to quantify anterior rectocele and to test whether this could substantiate clinical decision-making for operative treatment for anterior rectocele **Inclusion criteria:** Women with suspected ARC **Exclusion criteria:** Not described

Dellemare 1994 *(Continued)*

Patient characteristics and setting	**Nr of included patients:** 33 **Gender:** Female (100%) **Age:** mean age 57 years, range 25 - 78 years **Symptoms:** ARC-related symptoms: a feeling of outlet obstruction, the need to support the anterior rectal wall during evacuation, false urgency, faecal incontinence (also during coitus), and incomplete evacuation on defaecation **Ethnicity:** Unknown **Co-morbidities:** Unknown **Setting:** Tertiary care, single centre **Time period:** April 1990 and August 1992 **Country study is conducted:** The Netherlands
Index tests	**Name index test:** Dynamic MRI **Details of conducting index test:** Dynamic MRI was performed on a Philips 1.5 Tesla Gyroscan (Philips Medical Systems, Best, The Netherlands) without preparations like contrast introduction, diet, enema, or any other manipulation. The participants were examined in prone position, enabling air to collect in the rectum, creating an excellent natural contrast medium for MRI. The participants were asked to void before the examination, to prevent compression of the rectum by a full urinary bladder **Imaging acquisition:** 5 transverse slices through the pelvis were obtained, with the participants in prone position: spin echo, repetition time 350 milliseconds, echo time 20 milliseconds (spin echo 350/20), 10-mm slice thickness, 256 x 204 matrix, 1 signal acquisition, resulting in a 1-minute acquisition time sequence. One of the transverse views was used to plan the sagittal slices, using the same pulse sequence as in the transverse view. The midsagittal level was selected and this plane was used to obtain the dynamic images. The dynamic scan was performed with the gradient echo pulse sequence of 10 seconds acquisition time. The gradient echo sequence used a repetition time of 60 milliseconds, a flip angle of 60 ~ an echo time of 14 milliseconds, a field of view 30 x 30 cm, 1 signal acquisition, and a 256 x 154 acquisition matrix, resulting in an acquisition time of 10 seconds. The participants were instructed to suspend breathing during this period. 1 image was obtained at rest and 1 during maximal pelvic strain, resulting in 28 images. The specific instructions were in conformity with the instructions for radiographic defaecography **Imaging analysis:** The anorectal junction was defined in conformity with the radiographic defaecography as the intersection point of the central axis of the anal canal and the line along the posterior wall of the distal rectum. Because the tuber ischiadicum is not visible on the midsagittal slice, the baseline for dynamic MRI was defined as the junction line between the cranial side of the symphysis pubis and the distal sacrum. The distance between the projection of the anorectal junction and the anterior rectal wall on the baseline was defined by us as the quantitative size of the ARC. This distance was measured independently by 2 observers (RHK and JBVMD) on each MRI examination and the inter-observer difference was determined. Qualitative grading of the ARC was carried out double-blind in conformity with the radiographic defaecography **Threshold test positivity:** Rectocele: present/absent; pelvic floor descent: present/absent
Target condition and reference standard(s)	**Name index test 'EP':** Radiographic defaecography **Details of conducting evacuation proctography:** For radiographic defaecography, 120 ml of high-density BaSO4 contrast medium were introduced into the rectum with the participant in the LDP, followed by thickened BaSO4 contrast medium up to capacity, usually approximately 250 ml. Thickening was achieved by adding Metamucil | (Marion Merrell Dow, Inc., Cincinnati, OH) to BaSO4 contrast with a specific gravity of 1.2 g/cm 3, in a volume ratio of 1:30, to attain faecal viscosity **Imaging acquisition:** Films and video recordings were taken in lateral projection at rest, while squeezing, during Valsalva's manoeuvre, during coughing and during defaecation, with the participant seated on a modified toilet seat, mounted on the footplate of a remote control stand as described by others

Dellemare 1994 *(Continued)*

	Imaging analysis: We defined the anorectal junction as the intersection point of the central axis of the anal canal and the line along the posterior wall of the distal rectum in conformity with others. The junction line between the caudal part of the tuber ischiadicum and the coccyx is used as baseline for the assessment of the position of the anterior rectal wall in relation to the anorectal junction in radiographic defaecography. The distance between the projection of the anorectal junction and the anterior rectal wall on the baseline is defined by us as the quantitative size of the ARC. This distance was measured independently by 2 observers (RHK and JBVMD) and corrected for geometric enlargement in each radiograph. The difference of these 2 measurements was analysed in order to determine the inter-observer difference. Qualitative grading of the ARC on these 66 radiographs was carried out, double-blind, by 2 experienced observers who have analysed well over 1000 radiographic dynamic defaecographic studies each (RHK and JBVMD) **Threshold test positivity:** Rectocele: present/absent; pelvic floor descent: present/absent
Flow and timing	**Enrolment and exclusions (+ reasons):** Total enrolled in the study: 33. The first 19 participants included in this study who underwent dynamic defaecography were excluded from dynamic MRI because the interval between the 2 examinations was considered unacceptably long. All 14 participants who received both dynamic defaecography and dynamic MRI were included in the analysis **Nr analysed:** 14 **Time interval (+ interventions) between index test and reference standard:** Dynamic MRI within 1 month after radiographic defaecography
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			

Dellemare 1994 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Dellemare 1994 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Faggian 2013

Study characteristics	
Patient Sampling	**Patient selection:** 614 women with symptoms related to pelvic floor dynamic dysfunctions were enrolled in a retrospective study **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** To assess the diagnostic tools available to define the imaging strategy in patients with pelvic floor dynamic dysfunctions and to investigate their abilities in the diagnosis of enterocele, elytrocele and edrocele **Inclusion criteria:** Patients with symptoms related to pelvic floor dynamic dysfunctions **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 614 **Gender:** Female **Age:** Mean age was 57.3 years **Symptoms:** Referral symptoms varied from constipation and obstructed defaecation to incontinence **Ethnicity:** Unknown **Co-morbidities:** Unknown **Setting:** Tertiary care, single centre **Time period:** January 2008 to May 2011 **Country study is conducted:** Italy
Index tests	**Name index test:** Supine entero-magnetic resonance (SE-MR) **Details of conducting index test:** All SE-MR imaging studies were performed on a 1.5 T closed magnet (Magnetom Symphony, Siemens, Germany). All participants were supine imaged with a body-phase-array receiver coil. To ensure an adequate bladder filling, all participants were invited to drink 500 - 700 ml of water 10 - 15 minutes before examination. The rectum and the vagina were filled with 200 ml and about 25 - 30 ml, respectively, of ultrasonographic gel

Faggian 2013 *(Continued)*

	Imaging acquisition: After an initial localiser in 3 different planes, the study protocol included the following MR imaging sequences: TSE T2-W axial (matrix, 181 x 256; slices, 25 mm; thickness, 5 mm; TR/TE, 845/11; flip angel 150 °) sequences, and functional dynamic sequences TRUFISP T2-W sagittal, during squeezing, straining, pushing and evacuation (matrix, 181 x 256; slices 1; thickness 8 mm; TR/TE, 3.75/1.6; flip angle, 80 °). The SE-MR images so obtained were then assembled in cineview in post-processing. The examination took about 30 minutes to complete **Imaging analysis:** Both examinations were analysed by 2 expert investigators (RG, BF) blinded against both the clinical data and the results of the other imaging technique **Threshold test positivity:** Enterocele: descent of small bowel loops, peritoneal fat or sigmoid colon into the rectogenital space above the superior portion of the vaginal dome
Target condition and reference standard(s)	**Name index test 'EP':** Entero-colpo-defaecography (ECD) **Details of conducting evacuation proctography:** No bowel preparation was used for ECD. In order to obtain small-bowel contrast, 1 hour before the examination, 200 mL of barium sulfate 60% p/v was administered to each participant. Through a catheter inserted in the bladder 400 cc of iodine contrast medium (Ultfavfsf, Bayer Schcrfng Pharrrfa. Berlin, Germany) was injected through urinary catheterisation until the participant felt a sensation of fullness. Afterwards, the participant was placed in left lateral recumbent decubitus position, in order to inject 200 cc of barium paste (Prontobario Esofago I 13%, barium paste, Bracco, Milan, Italy) introduced into the rectum. During injector removal, the anal canal was also contrasted. The vagina was contrasted with 25 ml of barium paste. The fluoroscopic table was then tilted upright 90 °, and the participant was placed seated on a radiolucent commode **Imaging acquisition:** An anterio-posterior radiograph was taken with the participant at rest; after that, 5 lateral radiographs were taken at rest and during the following phases: squeezing, abdominal straining, pushing, evacuating, and at rest after evacuation **Imaging analysis:** Both examinations were analysed by 2 expert investigators (RG, BF) blinded against either the clinical data or the results of the other imaging technique **Threshold test positivity:** Enterocele: descent of small bowel loops, peritoneal fat or sigmoid colon into the rectogenital space above the superior portion of the vaginal dome
Flow and timing	**Enrolment and exclusions (+ reasons):** ECD and SE-MR was performed in all participants **Nr analysed:** 614 **Time interval (+ interventions) between index test and reference standard:** SE-MR was performed after ECD in the same day
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Faggian 2013 *(Continued)*

Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern

Faggian 2013 *(Continued)*

DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Faucheron 2014

Study characteristics	
Patient Sampling	**Patient selection:** A prospective study of a single-centre cohort was carried out in which a standardised evaluation was used by experienced surgeons and radiologists for all consecutive patients who were finally operated on for posterior pelvic floor prolapse. 50 women entered the study and 17 other patients who had previously undergone surgery for pelvic prolapse during the same period were excluded **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The accuracy of dynamic cystocolpoproctography (DCP) and dynamic MRI were compared in diagnosing posterior pelvic floor disorders **Inclusion criteria:** Women with posterior pelvic floor prolapse **Exclusion criteria:** Previous surgery for pelvic prolapse
Patient characteristics and setting	**Nr of included patients:** 50 **Gender:** Female (100%) **Age:** The median age of the patients was 53 (range, 31 – 81) years **Symptoms:** The presenting symptoms were obstructed defaecation in 50 (100%) participants; anal incontinence or soiling in 17 (34%); raised transit time constipation in 14 (28%); dyspareunia in 10 (20%); pelvic pain or heaviness in 9 (18%); and pain on defaecation in 2 (4%). Thirty-seven (74%) participants also complained of a posterior colpocele **Ethnicity:** Unknown **Co-morbidities:** Only 1 participant was nulliparous; the median parity was 3 (range, 0 – 8). 24 (48%) participants had undergone hysterectomy. 11 (22%) participants had a body mass index of > 25 kg/m^2 **Setting:** Tertiary care, single centre **Time period:** 2010 - 2012 **Country study is conducted:** France
Index tests	**Name index test:** Functional pelvic MRI **Details of conducting index test:** MRI was performed, with the participant in the supine position, using a 1.5 Tesla superconductive unit and a circularly polarised (quadrature) body coil (INTERA; Philips Electronics,

Faucheron 2014 *(Continued)*

	Koninklijke, the Netherlands) by a radiologist (DC; 9 years experience) specialised in pelvic imaging in women, who also interpreted the MRI. The examination involved vaginal and rectal opacification with 20 and 120 ml of sonographic transmission gel. Before the examination began, the participant was instructed by the radiologist of the manoeuvres that would be required during imaging **Imaging acquisition:** Following a morphological analysis of the pelvic organs at rest, these manoeuvres consisted of contraction of the pelvic floor muscles, followed by relaxation, straining and rectal evacuation (onto waterproof padding placed beneath the buttocks) for the proctographic phase. The pulse sequences included T2-weighted turbo spin-echo sequences (TR range/TE, 3300 - 3800/90; matrix size, 196 9 256) and 1 acquisition in the axial, coronal and sagittal planes. The second phase was the dynamic analysis, including T2-weighted turbo spin-echo sequences (2D FFE balances, 90 scans of 0.7 s, 5-mm thickness) in the sagittal plane. A post evacuation phase was added at the end of the examination to study the behaviour of the empty rectum **Imaging analysis:** Functional pelvic MRI was analysed using specific measurements to determine the presence of full-thickness rectal prolapse, rectal intussusception, rectocele, vaginal vault prolapse and peritoneocele (including enterocele, sigmoidocele, hedrocele and epiplocele). The radiologist was blinded to all clinical data and performed the examination according to a 'pelvic floor disorder – prospective protocol'. For the purpose of the study, the radiologist was also blinded to the results of the other imaging technique. The findings were recorded on a standardised form **Threshold test positivity:** The pubosacrococcygeal line was used as the reference point for defining the site of prolapse and its extent, except for rectocele, which was diagnosed if the anterior margin of the rectal wall bulge was more than 3 cm anterior to a line drawn along the long axis of the anterior anal canal
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic cystocolpoproctography (DCP) **Details of conducting evacuation proctography:** The DCP technique was essentially that described by Mahieu 1984, slightly modified by Shorvon 1989. The main refinements to the technique for the proctographic part have already been described extensively. The DCP was performed by a radiologist (AD, with 25 years experience) who had carried out more than 2300 video dynamic defaecographies before starting the study. Laxatives or enema were not given before the examination. The bladder was catheterised with a Foley catheter (8 Fr) and water-soluble iodine contrast was instilled to a maximum of 200 ml or until the participant felt distension of the bladder. The vagina was opacified with a mixture of barium and vaginal gel, and the rectum was then filled with semisolid contrast material of standardised consistency composed of barium suspension mixed with starch, injected with a caulking gun injector until the participant felt rectal fullness **Imaging acquisition:** Lateral radiographs and a videotape with participants in the sitting position were obtained at rest, on squeeze and during and after evacuation. After rectal and bladder emptying at the end of the examination, further radiographs were taken with the participant straining maximally to show the full extent of the prolapse. The barium was weighed before injection into the rectum and after evacuation. The time for rectal evacuation was measured. The flow rate and post-defaecation residue were routinely calculated to exclude rectal akinesia **Imaging analysis:** DCP was analysed using specific measurements to determine the presence of full-thickness rectal prolapse, rectal intussusception, rectocele, vaginal vault prolapse and peritoneocele (including enterocele, sigmoidocele, hedrocele and epiplocele). The radiologist was blinded to all clinical data and performed the examination according to a 'pelvic floor disorder – prospective protocol'. For the purpose of the study, the radiologist was also blinded to the results of the other imaging technique. The findings were recorded on a standardised form **Threshold test positivity:** The pubosacrococcygeal line was used as the reference point for defining the site of prolapse and its extent, except for rectocele, which was diagnosed if the anterior margin of the rectal wall bulge was more than 3 cm anterior to a line drawn along the long axis of the anterior anal canal
Flow and timing	**Enrolment and exclusions (+ reasons):** DCP and functional pelvic MRI was performed in all participants **Nr analysed:** 50 **Time interval (+ interventions) between index test and reference standard:** Median of less than 1 week with maximum of 2 weeks
Comparative	

Faucheron 2014 *(Continued)*

Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test,			Low concern

Faucheron 2014 *(Continued)*

its conduct, or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Fiaschetti 2013

Study characteristics	
Patient Sampling	**Patient selection:** 49 consecutive patients who had symptoms of chronic constipation, feeling of incomplete evacuation, pain during defaecation, and/or faecal incontinence were enrolled. All the patients were referred by certified colorectal surgeons and underwent prior outpatient examinations, including digital examination and proctoscopy **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To assess the feasibility of magnetic resonance defaecography (MRD) in pelvic floor disorders using an open tilting magnet with a 0.25 T static field and to compare the results obtained from the same participant both in supine and orthostatic positions **Inclusion criteria:** Women with chronic constipation, feeling of incomplete evacuation, pain during defaecation, or faecal incontinence, or both **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 49 **Gender:** Female (100%) **Age:** Mean age 43.5 years, range 22 - 65 years **Symptoms:** Symptoms of chronic constipation, feeling of incomplete evacuation, pain during defaecation, or faecal incontinence, or both. Feeling of incomplete evacuation 35/49, Pain during defaecation 7/49, faecal incontinence 10/49, chronic constipation 41/49, sense of rectal bulging 18/49, dyspareunia 14/49, sense of vaginal bulging 10/49, feeling of incomplete urination 6/49, dysuria 10/49, sense of vesical bulging 3/49 **Ethnicity:** Unknown **Co-morbidities:** 1 participant had previously undergone stapled trans-anal rectal resection (STARR) for obstructed defaecation and rectocoele, and 4 others had undergone a hysterectomy to resect a fibroid uterus. 3 participants were nulliparous, with the remaining women having a mean of 1.3 children **Setting:** Tertiary care, single centre **Time period:** May 2010 - November 2011 **Country study is conducted:** Italy
Index tests	**Name index test:** Magnetic Resonance Defaecography (MRD) **Details of conducting index test:** MRD was performed using a permanent open magnet with changeable positions and static 0.25 T field, dynamic gradients with 20 mT/m power and 25 mT/m/s slew rate (G-SCAN, Esaote S.p.A., Genova, Italy). The magnet table was provided with a tilting mechanism from 0 ° to 90 ° with 2 ° steps, and allowed the evaluation both in supine and orthostatic positions. A surface lumbar spine DPA coil was used as the receiving coil, composed of a stiff base (length 320 mm x depth 280 mm x height 45 mm) and a flexible anterior band with variable dimensions (big band 89 x 18.5 cm; little band 69 x 18.5 cm) depending on the size of each participant. The protocol used was developed in a previous pilot study performed at the authors' institution. Before the examination, the rectum was filled with approximately 200 ml of suspension media (mashed potatoes) mixed with 1 ml paramagnetic contrast media gadobutrol (Gadovist 1 mol/l, Schering AG, Berlin, Germany). The bladder was also filled with 180 ml physiological solution mixed with 3 ml paramagnetic contrast media gadobutrol (Gadovist 1 mol/l) via a 16 F double-way Foley catheter, which remained in place during the entire study. Finally, the vagina was filled with an echographic gel suspension (Aquasonic 100, Parker Laboratories, Fairfield, NJ, USA) mixed with 0.5 ml paramagnetic contrast media gadobutrol. The mean time required to prepare the participant was 20 minutes (range 14 - 27 minutes) **Imaging acquisition:** Initially the examination was performed in the orthostatic position, with the magnet table pitched at 80 °. The 3 orthogonal image planes, were acquired at rest using the 3D HYCE sequence (hybrid contrast enhanced), a type of gradient echo balanced sequence with the following characteristics: 10 ms

Fiaschetti 2013 *(Continued)*

	repetition time (TR); 5 ms echo time (TE); 90 ° Flip Angle; 20 section; 2.5 mm section thickness; 280 x 280 field of view; 200 x 160 matrix. The static images were acquired in the sagittal plane at rest, and during sphincter contraction and straining using a GE T1-weighted sequence with the following parameters: 35 ms TR; 10 ms TE; 90[1]flip angle; 1 section; 5.5 mm section thickness; 300 x 300 FOV; 192 x 128 matrix. Finally, the dynamic phase was performed during defaecation using a GE T1-weighted sequence in the sagittal plane and with the following parameters: 30 ms TR; 6 ms TE; 90 ° flip angle; 1 section; 5.5 mm section thickness; 300 x 300 FOV; 192 x 128 matrix; 3 s/image acquisition time. T2-weighted sequences were not available. The same study protocol was followed in the supine position with the magnet table pitched at 0 °, prior to second filling of the rectal ampulla. Finally, the bladder catheter was removed and the urinary study was performed using the same dynamic sequences used in the defaecation study. The overall magnet time required to complete the acquisition in both positions and the second rectal filling was an average of 68 minutes (range 42 - 93 minutes) **Imaging analysis:** All examinations were evaluated separately by 2 radiologists to establish inter-observer concordance. Both observers were experienced in PFD study (3 years of experience for the first observer, 1 year of experience for the second) and they repeated the measurements 1 month later to evaluate intra-observer concordance **Threshold test positivity:** Rectocele: any; enterocele: > 1 cm below PCL; intussusception: full-thickness or mucosal; pelvic floor descent: ARJ more than 2 cm below PCL
Target condition and reference standard(s)	**Name index test 'EP':** Colpo-cysto-defaecography **Details of conducting evacuation proctography:** The colpo-cysto-defaecography was acquired on a remote-controlled digital radiological system OPERA T90cex (General Medical, Merate, Italy) in the sitting position through a dedicated radio-transparent device. The pelvic organs were prepared as follows: the vagina was filled with 50 ml echographic gel (Aquasonic 100, Parker Laboratories, Fairfield, NJ, USA) mixed with 5 ml iodinated contrast media (iopamidol, 370 mg iodine/ml; Iopamiro 370, Bracco S.p.A., Milano, Italy); the bladder was filled with 120 ml physiological solution and 120 ml iopamidol (Iopamiro 370) through a 16 F double-way bladder Foley catheter, left in situ; the rectum was filled with 180e240 ml barium paste (Prontobario 110% p/v, Bracco S.p.A, Milano, Italy); the ileal loops were filled with 250 ml oral barium solution (Prontobario HD, Bracco S.p.A., Milano, Italy), about 45 min before the examination **Imaging acquisition:** At first, static images were acquired in the lateral, anteroposterior, and oblique projections, respectively, at rest, and during contraction, and straining, and then we performed a dynamic defaecatory phase acquired only in the lateral projection using serial imaging (1e3 images/s). After removing the bladder catheter, dynamic images were acquired using serial imaging (1 - 3 images/s) in the oblique projection in order to evaluate the urinary function. The technical parameters were the following: 80e90 kV, 100 mAs and focus-sensitive plain distance 1.10 m. **Imaging analysis:** All examinations were evaluated separately by 2 radiologists to establish inter-observer concordance. Both observers were experienced in PFD study (3 years of experience for the first observer, 1 year of experience for the second) and they repeated the measurements 1 month later to evaluate intra-observer concordance **Threshold test positivity:** Rectocele: any; enterocele: > 1 cm below PCL; intussusception: full-thickness or mucosal; pelvic floor descent: ARJ more than 2 cm below PCL
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants enrolled were included in the 2 x 2 table **Nr analysed:** 49 **Time interval (+ interventions) between index test and reference standard:** Unknown
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Fiaschetti 2013 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			High
DOMAIN 3: Reference Standard			

Fiaschetti 2013 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Foti 2013

Study characteristics	
Patient Sampling	**Patient selection:** 19 consecutive patients were included in the study **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To prospectively compare the diagnostic capabilities of magnetic resonance (MR) imaging with conventional defaecography (CD) in outlet obstruction syndrome

Foti 2013 *(Continued)*

	Inclusion criteria: Patients with clinical symptoms of outlet obstruction **Exclusion criteria:** N/A
Patient characteristics and setting	**Nr of included patients:** 19 **Gender:** 2 men and 17 women (data on women only received from authors) **Age:** Mean age 54 years; range 36 – 77 years **Symptoms:** Outlet obstruction syndrome (100%) associated with urinary and faecal incontinence in 3/19 participants (15%) and with urinary difficulty in 1/19 participants (5%) **Ethnicity:** Not described **Co-morbidities:** 6 (35%) of the 17 women had undergone hysterectomy. Of the 19 participants, 9 (47%) had a history of pelvic surgery, which included anal fissure, staple transanal rectal resection (STARR), rectopexy, endometriosis, reconstructive pelvic floor surgery, cystopexy, anterior rectocele and haemorrhoidectomy. Of the 17 women, 5 were nulliparous and 12 had had 1 - 6 deliveries **Setting:** Secondary care, single centre **Time period:** Between July 2007 and January 2009 **Country study is conducted:** Italy
Index tests	**Name index test:** MRI **Details of conducting index test:** MR examinations were performed with a closed-configuration superconducting unit with a 1.5-T field strength (GESigna HDx 1.5 T, GE Medical Systems, Milwaukee, WI, USA) using an 8-channel torso coil. Participant preparation and co-operation are essential for the success of the study. Prior to the examination, participants are given an enema and instructed about the manoeuvres to be performed inside the magnet. Participants are invited to wear a large pad, a stratagem that has the dual purpose of preventing soiling of the MR bed and reducing psychological discomfort. The bladder should be half full. Inside the gantry, the rectum is distended with approximately 150 ml of ultrasound gel (hyperintense on T2 and FIESTA sequences) introduced through a Nelaton catheter (20 Ch, 6.67 mm×360 mm) (Bicakcilar, Istanbul, Turkey) and a 50-ml catheter-tip syringe. The degree of straining is monitored with a respiratory gating device placed around the participant's waist. Inside the gantry, the participant lies supine (feet first), with knees slightly flexed, as this position facilitates evacuation of rectal contrast agent during defaecation **Imaging acquisition:** Our protocol includes the acquisition of: – High-spatial-resolution static sequences to study the morphology of the levator ani; – Dynamic sequences to study abnormalities of the pelvic organs during contraction, rest, straining and defaecation Static sequences included T2-weighted fast spin-echo (FSE) sequences in the sagittal, axial and coronal planes The technical parameters for this sequence were time to repetition (TR)/time to echo (TE), 4675/100; flip angle, 90 °; section thickness, 4 mm; interslice gap, 1 mm; bandwidth, 41.67 kHz; field of view (FOV), 32 cm; matrix, 320×224; number of averages, 4; number of images, 26; acquisition time, 3 minutes 49 seconds. Dynamic sequences were performed in the midsagittal plane identified on the T2-weighted FSE static images, with the pubic symphysis, urethra, vagina, rectum and coccyx included in the FOV. In the dynamic phase, 2 types of sequences were used: T2-weighted single-shot fast spinecho (SSFSE) and fast imaging employing steady-state acquisition (FIESTA) sequences acquired with the following parameters: – SSFSE (TR/TE, 708/90; flip angle, 90 °; section thickness, 8 mm; bandwidth, 83.3 kHz; FOV, 34 cm; matrix, 384 × 224; number of averages, 0.5; acquisition time for each image, 0.3 s) in the midsagittal plane, with sequential acquisition during contraction, rest and straining; – FIESTA (TR/TE, 3.3/1.4; flip angle, 45 °; section thickness, 8 mm; bandwidth, 125 kHz; FOV, 35 cm; matrix, 224 × 224; number of averages, 1; number of images, 20; acquisition time, 20 seconds) in the midsagittal plane, with continuous multiphase acquisition during contraction, rest, straining and defaecation. When clinical ex-

Foti 2013 *(Continued)*

	amination suggested the presence of lateral rectocele or lateral prolapse, the dynamic sequences were obtained in the axial and coronal planes as well. Overall examination time, including participant preparation, was approximately 40 minutes **Imaging analysis:** MR imaging was performed by a radiologist with 6 years' specific experience. The radiologists were aware of the results of the clinical examinations. The radiologist reading the MR images was blinded to the results of CD. MR images were displayed on a Picture Archiving and Communication System (PACS) screen, and cine loop presentation was used for dynamic sequences. Degree of prolapse measurement was performed by using electronic landmarks; each measurement was taken 3 times and expressed as a mean value **Threshold test positivity:** Rectocele > 2 cm depth; enterocele: small bowel below PCL; intussusception: intra-rectal or intra-anal invagination; anismus: ARA more acute during straining; pelvic floor descent; anorectal junction > 5 cm below the PCL during straining
Target condition and reference standard(s)	**Name index test 'EP':** Conventional Defaecography (CD) **Details of conducting evacuation proctography:** CD was performed by a radiologist with 10 years' experience. The participant received 250 ml barium orally, about 1.5 - 2 hours before the examination. The rectum was opacified using high-density Barium enema (150 - 200 ml). The participant was seated on a radiolucent commode **Imaging acquisition:** Images were acquired in lateral views during contraction, rest, straining and defaecation, including a final post-evacuation view **Imaging analysis:** The radiologist was aware of the results of the clinical examinations. CD was performed before MR imaging in all cases **Threshold test positivity:** Rectocele > 2 cm depth; enterocele: small bowel below PCL; intussusception: intra-rectal or intra-anal invagination; anismus: ARA more acute during straining; pelvic floor descent; anorectal junction > 5 cm below the PCL during straining
Flow and timing	**Enrolment and exclusions:** All participants were studied with pelvic CD (entero-defaecography) and MR imaging **Nr analysed:** 19 **Time interval (+ interventions) between index test and reference standard:** Not exceeding 1 month (mean 12 ± 4 days; range 4 – 26 days)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients		Low risk	

Foti 2013 *(Continued)*

have introduced bias?			
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its		Low risk	

Foti 2013 *(Continued)*

conduct, or its interpretation have introduced bias?			
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Grasso 2007

Study characteristics	
Patient Sampling	**Patient selection:** 43 women with either faecal incontinence or obstructive defaecation and no history of vaginal surgery or prolapse were referred to our diagnostic imaging department from the proctology outpatient service. All participants gave their informed written consent to undergo CCD and introital ultrasound examination **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To compare introital ultrasound with colpocystodefaecography (CCD) in quantifying the anorectal angle and in the diagnosis of posterior pelvic floor disorders **Inclusion criteria:** Women with functional impairment of the posterior pelvic floor **Exclusion criteria:** History of vaginal surgery or prolapse
Patient characteristics and setting	**Nr of included patients:** 43 **Gender:** Female 100% **Age:** The median age was 58 (range, 20 –79) years **Symptoms:** Either faecal incontinence or obstructive defaecation

Grasso 2007 *(Continued)*

	Ethnicity: Unknown **Co-morbidities:** The median body mass index of the population was 23.7 (range, 17.8 – 40.7). The median number of vaginal deliveries was 3 (range, 0 – 5); there were 6 nulliparous women **Setting:** Secondary care, single centre **Time period:** Between October 2004 and May 2005 **Country study is conducted:** Italy
Index tests	**Name index test:** Introital ultrasound **Details of conducting index test:** Introital ultrasound was performed using a Sonoline Antares(Siemens AG, Erlangen, Germany) ultrasound machine, equipped with a 6.2-MHz EC9-4 probe, with the participant in a semi-recumbent position (110 ° sitting angle) in a gynaecological chair, with legs flexed and opened. The examination took about 10 minutes. The probe was covered with a protection latex condom and was positioned near the hymeneal ring, on the posterior wall of the vulva, and oriented to acquire oblique axial images of the anorectal junction **Imaging acquisition:** 4 images were acquired: an axial view to measure sphincter and puborectalis thickness, and a sagittal view of the anorectal junction at rest, during squeezing and during sustained straining.The hypoechoic posterior wall of the rectum should be visualised first during squeezing, to obtain a reference point for the following scans, and then during other manoeuvres. The vertex of the ARA was identified as the point at which the maximum change in the posterior hypoechoic rectal wall was observed during squeezing. Finally, a sagittal cine-loop recording during squeezing and straining was made. The ARA was measured between the posterior wall of the rectum and the longitudinal axis of the anal canal. To ensure consistency, the sagittal reference plane was established by first finding the V-shaped puborectalis muscle in the oblique axial plane, then rotating the probe 90 ° clockwise, to place the anterior structures on the left side of the screen and of the images **Imaging analysis:** A third radiologist (RFG or SP) performed the introital ultrasound examination. The radiologist performing the ultrasound analysis was blinded to the CCD results **Threshold test positivity:** Rectocele: any; intussusception: any; anismus: straining/rest ratio ≤ 1
Target condition and reference standard(s)	**Name index test 'EP':** Colpo-cysto-defaecography (CCD) **Details of conducting evacuation proctography:** CCD is occasionally used for the diagnosis of pelvic floor diseases. We use it routinely in place of proctography, to evaluate accurately both rectal and urogenital structures. CCD is based on a triphasic approach for both bladder and rectum. Instilling contrast medium in the bladder and the rectum allows diagnosis of a reduction of co-ordinate sphincterial control. In participants who were continent, the bladder was filled with up to 250 mL of hydrosoluble contrast medium (Iobitridol 350 mgI/mL, Xenetix Guerbet, France) until the desire to micturate was felt. A mixture of about 20 mL of barium paste (ProntoBario113g/100 mL Bracco, Italy) and 20 mL of 1% lydocaine chloride (luan Molteni Farmaceutici, Italy), to reduce discomfort during examination, were instilled into the vagina. Approximately 120 – 200 mL of diluted barium suspension (60 g/100 mL) – the same as that used for barium enema examination – was introduced to show the sigmoid. Up to 200 mL barium paste (113 g/100 mL) was then injected into the rectum until the maximum tolerated capacity or total volume was reached. The surgeon still needs morphological information about the anorectal junction and related functionality in faecal incontinence, so a 12 G Foley catheter was placed in the rectum and the balloon distended to avoid loss of contrast medium during squeezing and coughing manoeuvres **Imaging acquisition:** CCD images were acquired with the participant in the horizontal lateral position at rest. Then, with the participant in the seated position on a special commode, lateral films were acquired at rest and during squeezing and straining manoeuvres. A coughing manoeuvre was performed to evaluate stress incontinence. The evacuation phase was added to study rectocele behaviour and the presence of intussusception; a post-evacuation X-ray image was obtained at maximal straining **Imaging analysis:** Proctographic images of CCD were assessed by 2 expert radiologists (SP or CCQ and/or MS) who were blinded to clinical complaints and to the ultrasound report; differences in assessment were resolved by consensus

Grasso 2007 *(Continued)*

	Threshold test positivity: Rectocele: any; intussusception: any; anismus: inability to evacuate ⅔ of the contrast medium within 30 seconds
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants enrolled received both Introital US and CCD and were all included in the 2 x 2 table **Nr analysed:** 43 **Time interval (+ interventions) between index test and reference standard:** Unknown
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	

Grasso 2007 *(Continued)*

If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity prespecified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Gufler 1999

Study characteristics	
Patient Sampling	**Patient selection:** 32 women with symptoms or physical findings or both, suggesting urinary incontinence or prolapse of pelvic organs were examined preoperatively with ultrafast dynamic MRI **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** Dynamic magnetic resonance imaging (MRI) using a single shot fast spin-echo technique was evaluated as a non-invasive alternative to cysto-urethography or colpo-cysto-rectography in women with pelvic organ prolapse or urinary incontinence, or both **Inclusion criteria:** Women with symptoms or physical findings or both, suggesting pelvic organ prolapse or urinary incontinence, or both **Exclusion criteria:** Not described
Patient characteristics and setting	**Nr of included patients:** 32 (12 colpo-cysto-rectography, 20 bead-chain cysto-urethrography) **Gender:** Female (100%) **Age:** mean 61 year, range 36 – 81 years **Symptoms:** Symptoms or physical findings or both, suggesting urinary incontinence or prolapse of pelvic organs **Ethnicity:** white **Co-morbidities:** 27 participants had 2+ children, 2 were nulliparous, and 3 had 1 child. 6 participants were premenopausal, 2 in the menopause, and 24 postmenopausal. Colpo-cysto-rectography was performed in 10 hysterectomised participants and in 2 without history of hysterectomy but with clinical suspicion of rectocystoceles. Bead-chain cystourethrography was performed in 20 participants who had not undergone hysterectomy **Setting:** Secondary, single centre **Time period:** 1994 - 1995 **Country study is conducted:** Germany
Index tests	**Name index test:** Dynamic MRI **Details of conducting index test:** All participants were studied on a superconductive 1.0 T Magnetom-Expert scanner (Siemens, Erlangen, Germany). No contrast agent was applied for either dynamic or static MRI **Imaging acquisition:** A body phased-array coil was used for data collection. A T1-weighted turbo-gradient-echo pilot scan was performed to localise the bladder neck. If this structure was not adequately localised, the coil was repositioned and the pilot acquisition repeated. The dynamic ultrafast images were then obtained in sagittal planes, positioned exactly through the bladder neck and the rectum. We used the HASTE sequence for fast dynamic imaging (TR 10.9 msec; TE 87 msec; number of excitations 1; field of view 320 3 280 mm; and matrix 240 3 256). Only half the k-space was needed for measurement (echo train length 128); the k-space was expanded with the half-Fourier method to 240 lines. Only 1 slice per excitation could be obtained with an acquisition time of 2 seconds. The slice thicknesses were 10 and 5 mm. At the following positions 1 slice was acquired: relaxed pelvic floor, pelvic floor contracted, relaxed pelvic floor, moderate pelvic strain, maximal strain, relaxed pelvic floor. Between each excitation there was an interval of 2 seconds. After a longer rest the whole cycle could be repeated if necessary. T2-weighted fast-spin-echo sequences (TR 4500 msec; TE 120 msec; acquisitions 2; slice thickness 5 mm; matrix 320 3 512; field of view 320 – 380 3 270 mm) in the sagittal and axial planes were additionally performed in all participants to image the whole pelvis **Imaging analysis:** The dynamic images were pictured on hard copy and analysed. The multiple images at different degrees of straining were formatted into a cinematic loop, allowing pseudokinematic represen-

Gufler 1999 *(Continued)*

	tation of pelvic floor changes. Measurements were done on hard copies with use of the internal scale. We used the pubococcygeal line as a reference line for the pelvic floor (Fig. 1d). The pubococcygeal line in the sagittal views was defined as the connection between the inferior border of the symphysis and the levator ani sling insertion at the coccygis bone *Additional information from authors:* Evaluation for MRI and CCR was performed separately, with participant names not visible on MRI **Threshold test positivity:** *Additional information from authors:* Rectocele: > 1 cm depth; enterocele: small bowel loops below PCL line
Target condition and reference standard(s)	**Name index test 'EP':** Colpo-cysto-rectography (CCR) **Details of conducting evacuation proctography:** After complete voiding through catherisation, the urinary bladder was refilled with 250 mL of a water-soluble contrast agent (Peritrast 300, Koehler Chemie, Alsbach, Germany), and a metal chain was placed in the urethra. A peritrast-soaked swab was introduced into the vagina, and the rectum was filled with 80 mL of a barium suspension **Imaging acquisition:** Lateral projection radiographs were performed in an upright position with relaxed pelvic floor and under maximal strain **Imaging analysis:** *Additional information from authors:* Analysis of MRI images and Rx were done separately with time intervals up to months **Threshold test positivity:** *Additional information from authors:* Rectocele: > 1 cm depth, Enterocele: small bowel loops below PCL line
Flow and timing	**Enrolment and exclusions (+ reasons):** In 12 participants, a colpo-cysto-rectography was performed because of suspicion of vaginal vault prolapse (n = 7), rectocele (n = 10), or bladder descent (n = 9). 20 participants without history of hysterectomy received only bead-chain urethrocystography, because physical examination yielded no suspicion of rectocele, enterocele, or vaginal prolapse **Nr analysed:** 12 (20 bead-chain urethrocystography in women with only symptoms of urinary incontinence were excluded for this review) **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* Between 1 day and 1 week
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		High risk	
Are the included patients only female or are test accuracy data provided for only female participants?			

Gufler 1999 *(Continued)*

Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the			Low concern

Gufler 1999 *(Continued)*

reference standard does not match the question?		
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Gufler 2004

Study characteristics	
Patient Sampling	**Patient selection:** Colpo-cysto-proctography and dynamic MRI of the pelvic floor were performed on 52 participants who had urinary incontinence with or without pelvic organ prolapse. In 7 of these participants colpo-cysto-proctography was carried out in both the supine and upright positions **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** To test whether there are statistically significant differences between measurement results on colpo-cysto-proctography in the upright and the supine positions, and to correlate these results with dynamic MRI of the pelvic floor in the supine position **Inclusion criteria:** Women who had urinary incontinence with or without pelvic organ prolapse **Exclusion criteria:** Patients with no CCP in the supine position
Patient characteristics and setting	**Nr of included patients:** 7 **Gender:** Female (100%) **Age:** 57 years (range 48 – 68 years) **Symptoms:** Stress urinary incontinence was present in 7 participants and prolapse symptoms in 4 **Ethnicity:** white **Co-morbidities:** They were all postmenopausal; 2 of them were nulliparous **Setting:** Secondary care, single centre **Time period:** 2000

Gufler 2004 *(Continued)*

	Country study is conducted: Germany
Index tests	**Name index test:** Dynamic MRI **Details of conducting index test:** MR imaging was carried out on a 1.0 T Magnetom-Expert scanner (Siemens, Erlangen, Germany) with a body phasedarray coil for data collection **Imaging acquisition:** Dynamic ultrafast images were obtained in sagittal planes using a single-shot RARE sequence with half Fourier data acquisition (TEeff 87 ms; turbo factor 256; field-of-view 320 × 280 mm; and matrix 240 × 256, acquisition time 1 second, slice thickness 10 mm). Acquisitions were obtained with relaxed pelvic floor and at maximal pelvic strain **Imaging analysis:** *Additional information from authors:* Evaluation for MRI and CCR was performed separately, with participant names not visible on MRI **Threshold test positivity:** *Additional information from authors:* Rectocele: >1 cm depth; enterocele: small bowel loops below PCL line
Target condition and reference standard(s)	**Name index test 'EP':** Colpo-cysto-proctography (CCP) **Details of conducting evacuation proctography:** CCP was performed after the urinary bladder had been filled with 250 ml of water-soluble contrast agent, the urethra marked with a Hodgkinson's beaded chain, the vagina outlined with a swab soaked in contrast agent, and the rectum filled with barium suspension **Imaging acquisition:** Lateral projection radiographs were taken in the upright (standing) and supine position with relaxed pelvic floor and under maximal pelvic strain **Imaging analysis:** *Additional information from authors:* Analysis of MRI images and Rx were done separately with time intervals up to months **Threshold test positivity:** *Additional information from authors:* Rectocele: > 1 cm depth; enterocele: small bowel loops below PCL line
Flow and timing	**Enrolment and exclusions (+ reasons):** All 7 participants with CCP in supine and upright position were included in the 2 x 2 table **Nr analysed:** 7 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* Few hours to 5 days
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	

Gufler 2004 *(Continued)*

Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		

Gufler 2004 *(Continued)*

Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Hainsworth 2016

Study characteristics	
Patient Sampling	**Patient selection:** Consecutive women undergoing integrated total pelvic floor ultrasound and defaecation proctography for pelvic floor defaecatory dysfunction between 2011 and 2014 **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** This study assesses the accuracy of integrated total pelvic floor ultrasound for the detection of rectocele, intussusception, enterocele and dyssynergy compared with defaecation proctography as the gold standard in the assessment of women with pelvic floor defaecatory dysfunction **Inclusion criteria:** Women with pelvic floor defaecatory dysfunction who underwent both integrated total pelvic floor ultrasound and defaecation proctography **Exclusion criteria:** None
Patient characteristics and setting	**Nr of included patients:** 393 **Gender:** Female (100%) **Age:** Mean age 54.5 years (range 21 - 91) **Symptoms:** Pelvic floor defaecatory dysfunction (ODS and faecal incontinence) **Ethnicity:** *Additional information from authors received:* Mixed **Co-morbidities:** Parity (mean 2.3, median 2, range 0 – 10). Mode of delivery: At least 1 vaginal delivery 252; Nulliparous 29; Unknown 27; Caesarean section only 15. Previous pelvic surgery 128; Nil 195; Hysterectomy 90; TVT/TVTO 17; Anterior posterior vaginal repair 16; Colposuspension 5; Perineal approach to rectal, Prolapse repair 5; Rectopexy 3; Other (STARR, rectocele repair, hysteropexy) 4 **Setting:** Tertiary care, single centre **Time period:** May 2011 and November 2014 **Country study is conducted:** United Kingdom
Index tests	**Name index test 1:** Transperineal Ultrasound (TPUS) **Details of conducting index test:** Transperineal scanning was performed with the participant supine with the legs flexed (no bowel preparation, enema or contrast was used). TPUS was performed using a conventional curved array probe (6 MHz, field of view 70 °) rested on the perineum to gain dynamic 2-dimensional midplane sagittal views **Imaging acquisition:** The participant was asked to squeeze up, bear down and cough during each scan (digitally recorded) **Imaging analysis:** Dynamic transvaginal, transperineal and defaecation proctography images were retrospectively and independently reviewed by a blinded clinician

Hainsworth 2016 *(Continued)*

	Threshold test positivity: Rectocele: ≥ 2 cm; enterocele: small bowel least into the upper third of the vagina; anismus: failure to relax or a paradoxical increase in the ARA on straining **Name index test 2:** Transvaginal Ultrasound **Details of conducting index test:** Transvaginal scanning was performed with the participant supine with the legs flexed (no bowel preparation, enema or contrast was used). Transvaginal scanning was performed using a linear array endoscopic probe (12 MHz) to obtain dynamic 2-dimensional posterior and anterior mid-sagittal views **Imaging acquisition:** The participant was asked to squeeze up, bear down and cough during each scan (digitally recorded) **Imaging analysis:** Dynamic transvaginal, transperineal and defaecation proctography images were retrospectively and independently reviewed by a blinded clinician **Threshold test positivity:** Rectocele: protrusion of the anterior rectal wall over the perineal body; enterocele: presence of bowel between the rectum and vaginal wall; Intussusception: full-thickness circumferential; anismus: failure to relax or a paradoxical increase in the anorectal angle on bearing down
Target condition and reference standard(s)	**Name index test 'EP':** Defaecation proctography **Details of conducting evacuation proctography:** Defaecation proctography was performed with oral contrast (20 ml Gastrografin®, 100 ml Baritop®100 and 400 ml water 30 minutes before the procedure) and 120 ml of rectal paste (mixture comprising 1 sachet (200 g) of Baritop® Plus, 100 g Readybrek® and 300 ml warm water) **Imaging acquisition:** The participant sat on a commode between a c-arm to empty the rectum **Imaging analysis:** Dynamic transvaginal, transperineal and defaecation proctography images were retrospectively and independently reviewed by a blinded clinician **Threshold test positivity:** Rectocele: a bulge of the anterior rectal wall beyond the projected anterior rectal wall of ≥ 2 cm; enterocele: descent of contrast-filled small bowel loops onto the rectum to touch the rectal wall; intussusception: full-thickness circumferential; anismus: failure to relax or a paradoxical increase in the anorectal angle during attempted evacuation
Flow and timing	**Enrolment and exclusions (+ reasons):** 323 women were included in the 2 x 2 table. 70 were excluded due to images unavailable (24), images incomplete (38), poor-quality images (5) or tests performed over 2 months apart (3) **Nr analysed:** 323 **Time interval (+ interventions) between index test and reference standard:** Maximum of 2 months
Comparative	Not applicable
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		

Hainsworth 2016 *(Continued)*

Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as			Low concern

Hainsworth 2016 *(Continued)*

defined by the reference standard does not match the question?		
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Halligan 1996

Study characteristics	
Patient Sampling	**Patient selection:** 17 adult women, referred for evacuation proctography because of possible enterocoele **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** We describe a simple ultrasound technique to diagnose enterocele, which has been validated by comparison with proctography **Inclusion criteria:** Women with possible enterocele **Exclusion criteria:** Not described
Patient characteristics and setting	**Nr of included patients:** 17 **Gender:** Female 100% **Age:** Median age was 53 years, with a range of 25 - 66 years. **Symptoms:** Not described **Ethnicity:** Not described **Co-morbidities:** Not described **Setting:** Secondary care, single centre **Time period:** Not described **Country study is conducted:** United Kingdom
Index tests	**Name index test:** Vaginal endosonography **Details of conducting index test:** 2 glycerine suppositories were inserted into the rectum and retained for 15 minutes. The participants were then asked to empty the rectum. With the participant in the left lateral position a Bruel and Kjaer type 3535 ultrasound scanner fitted with a 1850

Halligan 1996 *(Continued)*

	rectal endoprobe (B & K Medical UK, Bracknell, UK) was placed into the vagina. The water-filled balloon was distended to establish contact. The probe was then manipulated so that the distal few centimetres of rectum, just above the anorectal junction, was visualised in cross-section posteriorly **Imaging acquisition:** The 10 MHz transducer was mechanically rotated to give a 360 ° cross-sectional image. The participant was then asked to bear down in order to precipitate any enterocoele. This was repeated at least 3 times to confirm the findings **Imaging analysis:** Not described **Threshold test positivity:** A diagnosis of enterocoele was made if the rectum became obscured by bowel loops during straining
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography:** 2 glycerine suppositories were inserted into the rectum and retained for 15 minutes. The participants were then asked to empty the rectum. The participant was then escorted to the fluoroscopy suite and EP performed using a standard technique. With the participant in the left lateral position the rectum was filled with 120 ml of barium paste (E-Z-Paste®, E-Z-EM, Co., Westbury, NY), via a bladder syringe. A gauze swab coated with barium paste was inserted into the vagina so that its apex lay at the level of the vaginal vault. The participant was then turned supine and the fluoroscopy table raised to the vertical position. The participant was next asked to step off while a specially-designed commode, containing 4 mm copper filtration, was placed on the footrest **Imaging acquisition:** The participant, while sitting upon the commode, was instructed to empty the rectum as quickly and completely as possible during lateral videofluoroscopy of rectal voiding. The examination was saved on videotape for analysis **Imaging analysis:** Not described **Threshold test positivity:** A diagnosis of enterocoele was made if the vaginal marker was displaced away from the anterior rectal wall during evacuation
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants received both vaginal endosonography and evacuation proctography and were included in the 2 x 2 table **Nr analysed:** 17 **Time interval (+ interventions) between index test and reference standard:** Both investigations were performed on the same day
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		

Halligan 1996 *(Continued)*

Could the selection of patients have introduced bias?	Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?		
Do the included patients only have ODS symptoms?		
Are there concerns that the included patients and setting do not match the review question?		Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)		
Was the threshold for test positivity pre-specified?	Yes	
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear	
Could the conduct or interpretation of the index test have introduced bias?	Unclear risk	
If a reference line was used, was it the PCL?		
For MRI was a scanner used with Tesla 1 or higher?		
Are there concerns that the index test, its conduct, or interpretation differ from the review question?		Low concern
DOMAIN 3: Reference Standard		
Was the threshold for test positivity pre-specified?	Yes	
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear	
Could the reference standard, its conduct, or its interpretation have introduced bias?	Unclear risk	
If a reference line was used, was it the PCL?		
Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		

Halligan 1996 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Healy 1997

Study characteristics	
Patient Sampling	**Patient selection:** 10 women with difficulty defaecating were examined with both evacuation proctography and dynamic MR imaging. To simplify the evaluation of measurement agreement between the evacuation proctography and dynamic MR imaging, we excluded women in whom evacuation proctography had shown intussusception and rectal prolapse **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of this study was to determine the agreement between measurements of the anorectal configuration made with dynamic MR imaging and with evacuation proctography **Inclusion criteria:** Women with difficulty defaecating **Exclusion criteria:** Women in whom evacuation proctography had shown intussusception and rectal prolapse
Patient characteristics and setting	**Nr of included patients:** 10 **Gender:** Female (100%) **Age:** median 61 years (range 38 - 82 years) **Symptoms:** Difficulty defaecating 100% **Ethnicity:** Unknown **Co-morbidities:** Unknown **Setting:** Tertiary care, single centre **Time period:** Unknown **Country study is conducted:** United Kingdom
Index tests	**Name index test:** Dynamic MR imaging **Details of conducting index test:** Dynamic MR imaging was performed within I month of evacuation proctography using a l.5-T superconducting magnet system (Signa: General Electric Medical Systems, Milwaukee, WI). No rectal preparation was used before scanning. Soft rubber tubes, 5 mm in diameter, were placed in the vagina and rectum to act as luminal markers. The participant lay supine on the table with knees flexed over a rubber support. Maximal straining down in suspended expiration was practiced before scanning, with the participant encouraged

Healy 1997 *(Continued)*

	to bear down as though emptying the bowels. Waterproof pads were placed beneath the participant to prevent embarrassment from leakage of urine or faeces **Imaging acquisition:** Images were obtained at rest and during maximal straining in the sagittal plane using a fast spoiled gradient-recalled acquisition in the steady state (GRASS) pulse sequence (flip angle. 30#{176}: TRITE, I I.6/4.2: field of view, 34 cm: slice thickness, 7 mm; interslice gap, 2 mm: niatrix size. 256 x 128; and 2 excitations). This sequence gave 10 slices in 31 seconds **Imaging analysis:** Not described **Threshold test positivity:** Rectocele: any depth; anismus; more acute ARA during straining
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography:** Evacuation proctography was performed first in all participants. A standard technique was used. 2 glycerin suppositories were administered and the participants were asked to empty the rectum. 120 milliliters of barium paste (E-Z-paste: E-Z-EM. Westbury. NY) was instilled rectally using a bladder syringe **Imaging acquisition:** The participants then sat on a specially-designed commode for videofluoroscopy that was undertaken in the lateral projection at rest and during rectal voiding **Imaging analysis:** Not described **Threshold test positivity:** Rectocele: any depth; anismus; more acute ARA during straining
Flow and timing	**Enrolment and exclusions (+ reasons):** All 10 women had dynamic MRI and EP and were included in the 2 x 2 table **Nr analysed:** 10 **Time interval (+ interventions) between index test and reference standard:** Within 1 month
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern

Healy 1997 *(Continued)*

DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Karaus 2000

Study characteristics	
Patient Sampling	**Patient selection:** 17 consecutive women outpatients with long-standing symptoms of anorectal obstruction were prospectively investigated. In all participants a digital examination and a proctoscopy was performed first, followed by anorectal endoluminal ultrasound and defaecography **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The objective of this study was to evaluate dynamic anorectal endosonography in diagnosis of enteroceles and to compare this technique with defaecography as a reference method **Inclusion criteria:** Female patients with long-standing symptoms of anorectal obstruction **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 17 **Gender:** Female (100%) **Age:** Mean age of 65 ± 11 years **Symptoms:** Their anorectal complaints and signs of anorectal obstruction were feeling of anal blockade in 14 (82%), feeling of incomplete evacuation in 16 (94%), prolonged defaecation in 15 (88%), constipation in 11 (65%), and frequent digital evacuation in 9 (53%) **Ethnicity:** Unknown **Co-morbidities:** 10 of 17 patients had a previous hysterectomy **Setting:** Tertiary care, single centre **Time period:** Unknown **Country study is conducted:** Germany
Index tests	**Name index test:** Dynamic**a**norectal endosonography **Details of conducting index test:** Anorectal endosonography was performed using a combination scanner with a transversal sector scanner and a sagittal curved array scanner (Kontron Instruments, 65 MHz ER-BI-T, 7,5 MHZ ER-BI-S, Neufahrn, Germany). The endosonographic probe was connected to the conventional ultrasound equipment (Kontron Instruments, AI 52000S). For the whole examination the participant was lying in the left lateral position. The rectum was emptied first using a saline enema **Imaging acquisition:** After conventional inspection of the rectal wall and the anal canal for neoplasms and sphincter abnormalities, a dynamic study was performed as follows: the sagittal curved array scanner was directed to the ventral rectal wall. The inner verge of the anal canal defined as the beginning of the thickening of the muscularis propria was localised. The minimal distance between the peritoneal cavity and the inner anal verge was determined during rest. This was defined as the minimal peritoneal-anal distance (PAD). The participant was then asked to strain maximally while the examiner gave way with the rigid endosonography rod according to the pelvic descent. This was not so effective in the supine position. Thus, the left-lateral position was used in all investigations. The PAD was determined again during maximal strain. Thereafter, the investigation was repeated. The respective smallest numbers of each PAD determination were used for analysis **Imaging analysis:** For each participant the difference between PAD at rest and during maximal strain was calculated (delta PAD). Reproducibility within participants was guaranteed by only 1 observer (MK) repeating the entire procedure twice, taking images for measurement at rest and at the point of maximal movement during straining **Threshold test positivity:** Unknown

Karaus 2000 *(Continued)*

Target condition and reference standard(s)	**Name index test 'EP':** Defaecography **Details of conducting evacuation proctography:** For defaecography, 200 ml of contrast medium was given so that the distal colon and the rectum were visualised. Small-bowel contrast was additionally given in participants where enteroceles were supposed **Imaging acquisition:** Unknown **Imaging analysis:** Unknown **Threshold test positivity:** Unknown
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants underwent endoluminal ultrasound. In 14 of 17 participants a defaecography was performed. In the remaining 3 participants proctoscopy revealed a mucosal prolapse, where anorectal endosonography was normal, and the participants decided to be treated locally by ligation therapy and refused additional defaecography at this time **Nr analysed:** 14 **Time interval (+ interventions) between index test and reference standard:** After the first investigation the participant was given a rest of at least 2 hours before the second investigation started
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Unclear		
Where the index test results interpreted without knowl-	Unclear		

Karaus 2000 *(Continued)*

edge of the results of the other index test(s)?			
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Unclear		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Unclear
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Kelvin 2000

Study characteristics	
Patient Sampling	**Patient selection:** 10 women with symptoms of prolapse and pelvic floor dysfunction were referred for dynamic MR cysto-colpo-proctography and dynamic fluoroscopic cysto-colpo-proctography by urogynaecologists at our institution **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The purpose of this study was to compare dynamic MR cysto-colpo-proctography with fluoroscopic cysto-colpo-proctography for both the detection and measurement of the extent of pelvic organ prolapse **Inclusion criteria:** Women with symptoms of prolapse and pelvic floor dysfunction **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 10 **Gender:** Female (100%) **Age:** Mean 65 years (range 44 – 79 years) **Symptoms:** Symptoms of prolapse and pelvic floor dysfunction **Ethnicity:** Unknown **Co-morbidities:** The mean parity of the participants was 3.5 (range, 3 – 5). All participants had a history of hysterectomy. Of 10 participants, 9 had undergone at least 1 pelvic floor reconstructive surgery other than hysterectomy; for 8 of these participants, the procedures were performed at other institutions **Setting:** Tertiary care, single centre **Time period:** During a 3-month period in 1999 **Country study is conducted:** USA
Index tests	**Name index test:** Dynamic MR cysto-colpo-proctography **Details of conducting index test:** The MR examination was performed in a similar manner to that described by Lienemann 1997, but with the important exception that the examination was separated into 3 phases: a cystographic phase, a proctographic phase, and a post-toilet phase. This triphasic approach mimicked the phases of dynamic fluoroscopic cysto-colpo-proctography. To make the examinations as identical as possible, the amount of contrast material introduced into each of the pelvic organs was the same for both the MR imaging and the fluoroscopic examinations. MR imaging was performed with the paticipant in the supine position with a 1.5-T superconductive unit and a circularly polarised (quadrature) body coil (Vision; Siemens, Erlangen, Germany). The participant was asked to empty her bladder on arrival at the department. Before the examination began, the participant was instructed about the voluntary manoeuvres to be performed during imaging. Manoeuvres consisted of progressive straining during the cystographic sequence, and a contraction of the pelvic floor muscles (squeezing) followed by relaxation, subsequent progressive straining, and rectal evacuation during the proctographic phase. The importance of rectal evacuation was emphasised to the participant; we explained that evacuation was essential to obtain complete information about the degree of prolapse. Waterproof padding was placed beneath the buttocks and thighs to limit participant embarrassment and to protect the table of the MR imaging unit. The participant's bladder was catheterised with a 12- French catheter, and 50 ml of isotonic saline solution was instilled **Imaging acquisition:** MR images were obtained at rest in the axial, sagittal, and coronal planes. The pulse sequences used at rest included T2-weighted turbospin-echo sequences (TR range/TE, 3300 – 3700/90; matrix size, 196 × 256; 1 acquisition; field of view, 270 – 350 mm; 5 mm thickness) in the axial, coronal, and sagittal planes. For the cystographic phase, the participant was asked to strain progressively while a dynamic series of images was obtained in the midsagittal plane using a true fast imaging in a steady-state free precession sequence (TR/TE, 6.32/3.00; flip angle, 70 °; matrix size, 192 × 256; field of view, 250 – 330 mm; 1 image every 1.2 seconds).

Kelvin 2000 *(Continued)*

	The participant's bladder was then drained through the catheter. At that time, it was sometimes necessary to perform manual reduction of a large cystocele to promote bladder emptying. The catheter was then withdrawn and the participant was asked to void in a bathroom before the proctographic phase of the examination The proctographic phase of the examination involved both vaginal and rectal opacification. The vagina and then the rectum were opacified with 20 ml and 200 ml, respectively, of sonographic transmission gel (Aquasonic 100; Parker Laboratories, Fairfield, NJ) introduced through a 26-French catheter. The participant was asked to perform the rest-squeeze-relax-strain-evacuate manoeuvre. During this process, a dynamic series of images was obtained in the midsagittal plane using a true fast imaging in a steady-state free precession sequence. The rest-squeeze-relax-strain-evacuate manoeuvre and the imaging were repeated so that imaging during complete rectal evacuation could be obtained. The participant was asked to strain while a brief dynamic series of images was obtained in the axial plane at the level of the pubic symphysis using the true fast imaging in a steady-state free precession sequence. The axial plane and pubic location were chosen to assess the presence of pelvic floor ballooning. The participant then went into the bathroom again and was asked to attempt further rectal evacuation. On return from the bathroom, the post-toilet phase was performed to evaluate for enterocele, sigmoidocele, or peritoneocele. A second dynamic series was obtained in the coronal plane through the posterior pelvis to assess the extent of levator ani muscle descent. All the dynamic series were shown in cineloop presentation and recorded on videotape **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: any; enterocele: any
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic fluoroscopic cysto-colpo-proctography **Details of conducting evacuation proctography and imaging acquisition:** The technique of fluoroscopic dynamic cysto-colpo-proctography performed on a commode in the sitting position was similar to the technique previously described. Several modifications were instituted for this study. First, a preliminary radiograph (36 × 43 cm) was obtained to identify the pubococcygeal line, which extends from the inferior margin of the pubic symphysis to the sacrococcygeal joint. Second, only 50 ml of contrast material was used to fill the bladder because cystocele size is not affected by introducing larger volumes of contrast material. Third, a post-toilet image with maximal strain was routinely used because further rectal evacuation maximises the visualisation of enteroceles and sigmoidoceles. Preparation for the fluoroscopic examination required that the participant ingest 500 ml of barium suspension to opacify the small bowel. A preliminary radiograph was obtained, the bladder was catheterised, and 50 ml of diatrizoate sodium (Hypaque 50%; Winthrop Pharmaceuticals, New York, NY) was introduced. 2 lateral radiographs of the bladder were obtained, 1 at rest and the other on maximal strain. Bladder drainage was performed, the catheter was withdrawn, and the participant was asked to void in the bathroom The vagina was opacified with 20 ml of a mixture of barium and a vaginal gel (Acigel; Ortho Pharmaceutical, Raritan, NJ). A folded gauze square was inserted in the introitus to limit the loss of barium. The rectum was filled with 200 ml of a thick barium paste (Anatrast; Lafayette Pharmacal, Lafayette, IN). Lateral radiographs were obtained at rest, on squeeze, and during and after evacuation. The post-evacuation radiograph was obtained with the participant straining maximally, as was the post-toilet radiograph. The entire examination was recorded on videotape. Measurements of midline structures corrected for magnification were made possible by the incorporation of a midline radiopaque centimetre ruler within the commode **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: any; enterocele: any
Flow and timing	**Enrolment and exclusions (+ reasons):** All enrolled participants received both examinations and were all included in the 2 x 2 table **Nr analysed:** 10

Kelvin 2000 *(Continued)*

	Time interval (+ interventions) between index test and reference standard: The MR imaging study routinely preceded the fluoroscopy study. For each participant, the 2 examinations were performed within 2 weeks of each other
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	

Kelvin 2000 *(Continued)*

If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		

Kelvin 2000 *(Continued)*

Could the patient flow have introduced bias?	Low risk

Lienemann 1997

Study characteristics	
Patient Sampling	**Patient selection:** We examined 44 women with an isolated or combined descensus of the pelvic floor compartments and stress urinary incontinence **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The purpose of our study was to combine dynamic MRI and adequate opacification to better delineate the pelvic-floor anatomy and to visualise the extent of descensus and prolapse. We compared this technique to dynamic fluoroscopy (DF) using the clinical evaluation and the intraoperative results as reference **Inclusion criteria:** Women with isolated or combined visceral descent and stress urinary incontinence **Exclusion criteria:** Not described
Patient characteristics and setting	**Nr of included patients:** 44 **Gender:** Female 100% **Age:** Mean age 61 years, range 32 – 83 years **Symptoms:** Isolated or combined descensus of the pelvic floor compartments and stress urinary incontinence **Ethnicity:** Unknown **Co-morbidities:** mean parity 2; range of parity 0 – 6. 34 participants had previous hysterectomy (77%) and 84% had a history of prior reconstructive surgery for genital prolapse **Setting:** Tertiary care, single centre **Time period:** Unknown **Country study is conducted:** Germany
Index tests	**Name index test:** MRI **Details of conducting index test:** Magnetic resonance imaging was performed with a 1.5- T superconductive magnet unit (Vision, Siemens, Erlangen, Germany). The empty bladder was filled with 60 ml of isotone saline solution using a 26-F Foley catheter. In all cases the urethra was marked with a sterile cotton thread soaked with Magnevist (Schering AG, Berlin,nGermany). Opacification of the vagina was achieved either by using sonography gel or barium paste mixed with 50 ml of Magnevist enteral (4 cases). The rectum was filled with 200 ml sonography gel until the participant expressed an urge to relieve the bowels. The examination was performed with the participant lying head first, supine with the request to slightly open the legs. We used a body-array surface coil. With absorbent tissues we prevented running out. No premedication was used **Imaging acquisition:** Pulse sequences included T2-weighted turbo spinecho sequences (TR 3500–3800 ms, TE 99 ms, matrix 308 ´ 512, 2 acquisitions, field of view 370 – 250 mm) of the pelvis in axial, coronal, and sagittal orientation. With the exception of the sagittal orientation (3 mm), the slice thickness was 5 mm. For the dynamic examination, the thread in the urethra which was seen on the axial image was used as reference point for the

Lienemann 1997 *(Continued)*

	sagittal single-slice True-FISP sequence (TR 5.8 ms, TE 2.5 ms, flip angle 70 °, matrix 256 x 256, field of view 270 mm, a total of 30 measurements with 1 image every 1.2 seconds, in plane resolution 1.02 mm). During the examination the participants were asked to relax the pelvic-floor muscles, contract them slowly, and then relax again. Then the participant was asked to increase the intra-abdominal pressure by straining and then relaxing. This cycle was repeated twice to a maximum of 4 times. The dynamic imaging sequences were presented in a cine loop and videotaped. The overall time of examination varied between 20 and 30 minutes **Imaging analysis:** The interpretation of the images and cine loops from each modality was performed by 2 experienced radiologists who were blinded to the results of the other investigations. In case of disagreement a consensus was attempted **Threshold test positivity:** Rectocele > 3 cm, Enterocele: widening of the rectovaginal space or deepening of the pouch of Douglas
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic fluoroscopy **Details of conducting evacuation proctography:** Dynamic fluroscopy (Polystar II, Siemens, Erlangen, Germany) of all participants was performed in an upright position. Opacification included the bladder (water-soluble contrast media), the urethra (a thread soaked with water-soluble contrast media), the vagina (approximately 30 ml of barium paste (Micropaque, Guerbet, France) using a forceps with a soaked folded gauze square), and the rectum (approximately 200 ml of barium; Micropaque, Guerbet). No opacification of the pelvic small bowel was used. The sigmoid colon was not completely opacified routinely. For measurement of midline structures corrected for magnification a radiopaque ruler was taped to the participants rima ani **Imaging acquisition:** Fluoroscopic evaluation was performed by 1 of the authors (AB) and included a series of 1 image a second during squeezing and straining **Imaging analysis:** The interpretation of the images and cine loops from each modality was performed by 2 experienced radiologists who were blinded to the results of the other investigations. In case of disagreement a consensus was attempted **Threshold test positivity:** Rectocele > 3 cm; enterocele: widening of the rectovaginal space or deepening of the pouch of Douglas
Flow and timing	**Enrolment and exclusions (+ reasons):** MR-CCRG as well as in DF 2 participants were not able to squeeze and strain properly. Another participant suffered from claustrophobia and therefore MR-CCRG could not be performed. 1 DF was inconclusive because of a participant with superimposition by bilateral total hip prostheses. So results of 40 (out of 44) participants were included in the 2 x 2 tables **Nr analysed:** 40 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* Normally less than 7 days
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		

Lienemann 1997 *(Continued)*

Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity prespecified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity prespecified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		

Lienemann 1997 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Lienemann 2000

Study characteristics	
Patient Sampling	**Patient selection:** We examined 55 women. Gynaecologic examination confirmed an isolated or combined pelvic floor descent in all participants. **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of this study was to evaluate magnetic resonance colpo-cysto-rectography in the diagnosis of enteroceles **Inclusion criteria:** Women with isolated or combined pelvic floor descent **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 55 **Gender:** Female 100% **Age:** Mean age 61 years, range 32 - 84 years **Symptoms:** Symptoms of pelvic floor descent **Ethnicity:** Unknown

Lienemann 2000 *(Continued)*

	Co-morbidities: In their past surgical history 84% of participants had undergone at least 1 previous gynaecologic operation related to pelvic floor renovation, which in 77% of these cases included a hysterectomy **Setting:** Tertiary care, single centre **Time period:** Unknown **Country study is conducted:** Germany
Index tests	**Name index test:** MR Colpo-cysto-rectography **Details of conducting index test:** MRI was performed with a 1.5 Tesla System TM (Vision, Siemens Corp., Erlangen, Germany). All volunteers and participants were examined according to our MR-CCRG protocol previously described. Opacification of the bladder and urethra was omitted in all members of the control group, as was the urethral opacification in 41 of the participants. During MRI examination the partiicoant's position was supine with legs slightly apart, lying on an absorbent pad to cope with leakage **Imaging acquisition:** We used a body-array surface coil. During the examination with a true fast imaging with steady precession sequence (time of repetition 5.8 ms, time of echo 2.5 ms, Flip 70 °, Matrix 224 × 256, field of view 260 mm, pixel edge length 1.02 × 1.02 mm, slice thickness 7 mm; every cycle 30 single images with 1 picture/1.3 seconds), participants were asked, in synchrony with the pictures and starting from the relaxed position, to contract the pelvic floor muscles slowly and then to relax them again. Immediately afterward, the participant was asked to increase the intra-abdominal pressure progressively by straining before relaxing again. This cycle was repeated in sequence at least twice and up to four times, with the participant being encouraged to evacuate her bowels during the cycles. On completing the examination, the dynamic image sequences were arranged into an infinite loop and recorded on videotape. The procedure was performed by radiographers as part of the daily routine. The participant's preparation was done by the resident radiologist on duty and took approximately 2 minutes. The overall measurement time was 20 minutes **Imaging analysis:** 1 experienced radiologist (AL) and gynaecologist (CA) each performed blind evaluations of the gynaecologic and individual image sequence results obtained via the screening procedures. On disagreement, re-evaluation and consent were attempted **Threshold test positivity:** Enterocele: below PCL
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic cystoproctography **Details of conducting evacuation proctography:** All X-ray examinations were performed in a strict lateral projection using a Polystar II TM (Siemens Corp., Ertangen, Germany). Means for organ opacification are quoted in Table 1. A sterile feeding tube was introduced into the bladder. After drainage the bladder was filled with the contrast media. The tube was then replaced by a sterile cotton thread to outline the urethra. The vagina was opacified using a forceps with a soaked folded gauze square. Finally, after a digital rectal examination, the rectum was filled with the contrast medium. Opacification of the small bowel was not performed, and the sigmoid and descending colon also were not routinely filled with contrast material. No premedication was administered **Imaging acquisition:** Recording of data comprised sequences of 1 image per second. During examination of the pelvic floor, the standing participant was asked to contract and then to exert maximum strain once. Defaecation was explicitly encouraged. To prevent leakage, the participant was equipped with a sanitary towel and absorbent pads **Imaging analysis:** 1 experienced radiologist (AL) and gynaecologist (CA) each performed blind evaluations of the gynaecologic and individual image sequence results obtained via the screening procedures. On disagreement, re-evaluation and consent were attempted **Threshold test positivity:** Enterocele:**b**elow the pubococcygeal reference line and width of the rectovaginal space > 2 cm
Flow and timing	**Enrolment and exclusions (+ reasons):** For all participants further investigation consisted of MR-CCRG and for 34 of the 55 participants, an additional dynamic cystoproctography

Lienemann 2000 *(Continued)*

Nr analysed: 34

Time interval (+ interventions) between index test and reference standard: *Additional information from authors:* Less than 7 days

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation			Low concern

Lienemann 2000 *(Continued)*

differ from the review question?			
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	No		
Could the patient flow have introduced bias?		High risk	

Martellucci 2011

Study characteristics	
Patient Sampling	**Patient selection:** Between January and June 2009, all consecutive female patients with symptoms of ODS referred to a specialised coloproctology centre (General Surgery IV, Santa Chiara Hospital, Pisa) were evaluated prospectively. ODS was diagnosed according to the Rome II criteria, a clinical examination was performed in all women and information on bowel function, pregnancies, episiotomy, previous surgery and associated diseases was obtained. The severity of symptoms was assessed using the Cleveland Constipation Score (CCS). In all the patients a dynamic evacuation proctography (DEP) and a dynamic transperineal ultrasound (DTPU) were performed along with anorectal manometry and transanal ultrasound **Study design:** Cross-sectional test accuracy study, prospective

Martellucci 2011 *(Continued)*

	Study objective: The study aimed to compare DTPU, with defaecography and to evaluate its clinical value in patients with ODS. Furthermore, we aimed to determine whether DTPU can provide enough information to replace DEP. **Inclusion criteria:** Women with symptoms of ODS **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 54 **Gender:** Female (100%) **Age:** The median age of the participants was 59 (29 – 83) years **Symptoms:** Symptoms of ODS (100%) **Ethnicity:** Unknown **Co-morbidities:** The median parity with vaginal delivery was 2 (0 – 11) deliveries. 9 (16%) participants were nulliparous. 7 (13%) had previously undergone hysterectomy and 2 (3.7%) had previously had a STARR procedure **Setting:** Tertiary care, single centre **Time period:** January and June 2009 **Country study is conducted:** Italy
Index tests	**Name index test:** Dynamic Transperineal Ultrasound (DTPU) **Details of conducting index test:** DTPU was performed by an experienced investigator (MJ) blinded to the DEP results and to all the clinical data using a B&K Medical Pro Focus Scanner with a 6 MHz 8802 convex probe (B-K Medical, Herlev, Denmark). All the participants were examined after a rectal enema. With the participant in the lateral and gynaecological positions, the probe was placed on the perineum, applying very gentle pressure. The bladder was half filled and ultrasonographic gel was sometimes instilled into the rectum (15 cases) and the vagina (4 cases). **Imaging acquisition:** In all the participants, the anterior compartment (pubis, urethra and bladder), the middle compartment (vagina and rectovaginal septum) and the posterior compartment (anal canal, rectum and puborectalis muscle) were assessed. The images were obtained in longitudinal and transverse sections at rest and during straining. All the examinations were recorded **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele:10 mm depth, Enterocele: descent of intra-abdominal content on valsalva, Intussusception: Infolding of the rectal wall during straining, Anismus: ARA fails to open during straining
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic evacuation proctography (DEP) **Details of conducting evacuation proctography:** DEP was performed by 1 experienced investigator blinded to the DTPU results and to all clinical data, using the technique described by Kelvin 2000. The rectum was emptied by a glycerine suppository or enema. 1 hour before the examination, 300 ml of dilute barium suspension at 60% (Prontobario 60%; Bracco s.p.a., Milan, Italy) wase given orally to opacify the small bowel. Participants were asked to empty the bladder 5 minutes before filling the rectum with 200 ml of thick barium sulphate paste at 113% wt/vol with a consistency similar to that of faeces (Prontobario esofago; Bracco s.p.a.), injected with a syringe with the participant in the left lateral position. The vagina was opacified with 20 ml of barium paste. **Imaging acquisition:** Radiographs of the pelvis were then obtained at rest and on voluntary contraction of the pelvic floor muscles. The table was then moved into the upright position and the participant was seated on a commode for further exposures at rest and during squeezing. A left lateral view of the pelvis was taken during evacuation, with particular attention to the late evacuation phase. The entire exami-

Martellucci 2011 *(Continued)*

	nation was recorded on videotape and each videoclip was analysed using a computer video capture and digitising system combined with an image analysis programme **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: 10 mm depth, Enterocele: descent of intra-abdominal content on valsalva, Intussusception: circumfirential infolding of the rectal wall during straining, Anismus: ARA fails to open during straining
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants were included in the 2 x 2 table **Nr analysed:** 54 **Time interval (+ interventions) between index test and reference standard:** Maximum interval between tests was 3 months - *Additional data received from authors*
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the		Low risk	

Martellucci 2011 *(Continued)*

index test have introduced bias?			
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Martin 2017

Study characteristics			

Martin 2017 *(Continued)*

Patient Sampling	**Patient selection:** The selection process for participants was based on their initial ODS symptomatology for which they were referred to the colorectal surgery clinic, in a consecutive series from 2009 to 2012 **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of the present study was to evaluate the diagnostic accuracy of magnetic resonance (MR) defaecography and compare it with videodefaecography in the evaluation of obstructed defaecation syndrome **Inclusion criteria:** Patients were included in the analysis only if they met the Rome III criteria of functional constipation ODS **Exclusion criteria:** Patients who had other types of constipation, such as slow transit constipation or constipation due to irritable bowel syndrome, or those who refused testing were excluded. No other exclusion criteria were considered
Patient characteristics and setting	**Nr of included patients:** 40 **Gender:** 38 women / 2 men **Age:** Age mean: 59.5 (range: 35 - 79) **Symptoms:** Symptoms of ODS **Ethnicity:** Unknown **Co-morbidities:** The most common medical history was depressive syndrome, which was diagnosed in 8 (20%) cases. The median number of vaginal births was 2 (range 0 – 6). 23 of the women had previously undergone pelvic-abdominal surgery, and of these, 13 had undergone a hysterectomy **Setting:** Tertiary care, single centre **Time period:** Between 2009 and 2012 **Country study is conducted:** Spain
Index tests	**Name index test:** MR defaecography **Details of conducting index test:** The MR defaecography was carried out using a Siemens Magnetom Sonata closed MRI (Siemens Medical, Malvern, Pennsylvania, USA) of 1.5 Tesla (T). In preparation, an enema of 250 ml of water was administered 2 hours before the study. All of the participants were informed and instructed about the procedure. Participants were positioned lying on their back with their legs flexed on the resonance machine table. The rectal contrast material was prepared beforehand with 100 g of potato puree flakes, 400 g of barium sulfate, 7 ml of gadolinium, and water until a solution of 450 ml was reached. The contrast material was inserted into the rectum by means of a 50-ml syringe and was administered until a sensation of continual defaecation was achieved. No opacification procedure was carried out on the urinary bladder or vagina. Neither coils nor intravenous contrasts were used for the examination **Imaging acquisition:** To perform the midsagittal MR defaecography, high-definition video sequences were subsequently recorded in 3 dimensions using Fast Imaging with Steady State Precession (True FISP): at rest for 15 seconds, contraction for 15 seconds, and deaecation for 120 seconds or until complete defaecation was achieved. The sequences used to perform the magnetic resonance imaging were Axial T1W (In/Out phase), Coronal T2W, Sagittal True FISP, Sagittal Turboflash T1W (rest), Sagittal Turboflash T1W (contraction), and Sagittal Turboflash T1W (defaecation) **Imaging analysis:** MR defaecography was conducted by a single experienced radiologist. The assessors for videodefaecopgraphy (VD) and MR defaecography were blinded to the results of the other test and to the physical examination findings. After examinations were complete, all cases were discussed by a multidisciplinary committee on pelvic floor pathology **Threshold test positivity:** Rectocele: bulge extending more than 2 cm beyond the expected line of the anterior rectal wall; Enterocele: pelvic herniation during defaecation formed by an abnormally deep Douglas pouch contained by the small bowel, sigmoid colon or peritoneal fl (uid / mesenteric fat; intussuscep-

Martin 2017 *(Continued)*

	tion: descending full-thickness invagination of the rectal wall insufficient in descent to appear beyond the anal verge as an external rectal prolapse; anismus: thickening of the puborectalis muscle during prolonged evacuation of rectal contrast; pelvic floor descent: ARJ below PCL > 30 mm during defaecation
Target condition and reference standard(s)	**Name index test 'EP':** VD **Details of conducting evacuation proctography:** The VD was carried out with a conventional X-ray unit, high-resolution video recording equipment and a radiolucent seat of the lavatory bowl type. The participants were prepared by being informed and instructed about the procedure before the test was carried out. In addition, 1.5 hours before the test was carried out, we administered a cleansing enema and an oral solution of barium. Participants were asked to urinate immediately prior to the procedure. The participants were placed in the left lateral decubitus position with their legs flexed. The rectal contrast material, prepared with 200 g of potato puree flakes, liquid barium sulfate, and 700 ml of water, was progressively inserted into the rectum using a 60-ml syringe and a 4–6-mm lubricated catheter until a sensation of continual defaecation was achieved with 240 ml as the median amount (range 80 – 540 ml) **Imaging acquisition:** The participants were placed in the sitting posture on a radiolucent seat, and lateral fluoroscopy was performed during rest, contraction and defaecation **Imaging analysis:** VD was performed, analysed and evaluated by a gastroenterologist specialising in VD. The assessors for VD and MR defaecography were blinded to the results of the other test and to the physical examination findings. After examinations were complete, all cases were discussed by a multi-disciplinary committee on pelvic floor pathology **Threshold test positivity:** Rectocele: outpouching of the anterior rectal wall (any depth); enterocele: small bowel or sigmoid filling an abnormal peritoneal space in the pelvic floor; intussusception: rectum showing a funnel-shaped depression within the anal canal during push; anismus: the anorectal angle (ARA) unchanged during defaecation in comparison with the angle at rest; pelvic floor descent: difference of > 3.5 cm between the anorectal junction at straining and at rest
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants underwent both imaging tests (VD and MR defaecography) in the same order. In no cases were both tests performed on the same day **Nr analysed:** 40 **Time interval (+ interventions) between index test and reference standard:** The average time that elapsed between the completion of the first test (VD) and the second (MR defaecography) was 2 months (range 1 – 5 months). No other intervention was performed during this time interval
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	

Martin 2017 *(Continued)*

Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Was the threshold for test positivity pre-specified?			
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Martin 2017 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		High risk	

Matsuoka 2000

Study characteristics	
Patient Sampling	**Patient selection:** 14 consecutive women underwent surface coil MRI for the evaluation of faecal incontinence (5 women) or constipation (9 women) **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** This study assessed the value of common surface coil magnetic resonance imaging (MRI) in women with evacuatory disorders including faecal incontinence and constipation. In constipated participants the findings of videoprography and dynamic pelvic MRI were compared for the presence of rectocele, rectoanal intussusception, and sigmoidocele as well as the measurements of anorectal angle and perineal descent

Matsuoka 2000 *(Continued)*

	Inclusion criteria: Women with constipation or faecal incontinence **Exclusion criteria:** Women with complex symptoms or psychological disorders
Patient characteristics and setting	**Nr of included patients:** 14 **Gender:** Female (100%) **Age:** Mean age 59 years, range 40 – 78 **Symptoms:** Faecal incontinence (n = 5) or constipation (n = 9) (the 5 women who received imaging for the evaluation of fecal incontinence were excluded for this Cochrane review) **Ethnicity:** White **Co-morbidities:** Unknown **Setting:** Tertiary, single centre **Time period:** July 1996 to June 1997 **Country study is conducted:** USA
Index tests	**Name index test:** Dynamic pelvic MRI **Details of conducting index test:** Participants were positioned prone, and a 16-F urinary catheter inserted into the rectum, through which approximately 50 ml air was introduced as rectal contrast, and 10 ml air was used to inflate the balloon and fix the catheter **Imaging acquisition:** Breath-hold images were acquired in 19 seconds during standard proctography manoeuvres. T1-weighted sagittal and axial images were obtained through the pelvis at rest, while squeezing and during pushing manoeuvres. A Picker Vista Edge MRI (Picker, Highland Hights, Ohio, USA) and flexible extrabody coil were used for all evaluations **Imaging analysis:** In each case DPMRI was performed by a radiologist who was not made aware of the findings of VP until after his definitive report **Threshold test positivity:** Rectocele: ≥ 2 cm depth; Intussusception: any
Target condition and reference standard(s)	**Name index test 'EP':** Videoproctography (VP) **Details of conducting evacuation proctography:** VP was performed without any bowel preparation. The participants were placed in the left lateral decubitus position, after which, under fluoroscopic guidance, 50 ml liquid barium was introduced into the rectum. Following the liquid barium, up to 100 ml of thick barium paste, similar in consistency to stool, was injected also under fluoroscopic guidance into the rectum until the participant noted a sensation of rectal fullness (Anatrast EZM, Westbury, N.Y., USA). Injection was continued as the injector was withdrawn in order to outline the anal canal. The fluoroscopic table was then tilted upright 90 °, and the participant was seated on a water-filled radiolucent commode (Sunburst, Ladson, S.C., USA) **Imaging acquisition:** Lateral radiographics were obtained at rest and during squeezing and pushing; ultimately, participants were asked to evacuate. These processes were recorded on a videocassette tape using a high-resolution VHS recorder (model #AG6200, Panasonic, New York, N.Y., USA) **Imaging analysis:** VP was performed by surgeons in the Department of Colorectal Surgery and were all interpreted by a single surgeon who was kept unaware of the results of DPMRI until after a definitive VP report was issued **Threshold test positivity:** Rectocele: ≥ 2 cm depth; Intussusception: any
Flow and timing	**Enrolment and exclusions (+ reasons):** In participants with incontinence we compared the findings from endo-anal ultrasound (EAUS), anal MRI, and surgery for morphopathological findings of the internal and external anal sphincter components. In constipated participants the findings of videoproctography (VP) and dynamic pelvic MRI (DPMRI) were compared for the presence of rectocele, rec-

Matsuoka 2000 *(Continued)*

toanal intussusception, and sigmoidocele as well as the measurements of anorectal angle and perineal descent.

Nr analysed: 9 (the 5 women who received imaging for the evalufortion of faecal incontinence were excluded for this Cochrane review)

Time interval (+ interventions) between index test and reference standard: The examinations were performed at least within a month

Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct,			Low concern

Matsuoka 2000 *(Continued)*

or interpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Miravalle 2016

Study characteristics	
Patient Sampling	**Patient selection:** Women with symptoms of obstructed defaecation were included in the study. They performed defaecography and echodefaecography **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** This study was designed to validate the effectiveness of echodefaecography compared with defaecography in the assessment of anorectal dysfunctions related to obstructed defaecation **Inclusion criteria:** Women with symptoms of obstructed defaecation

Miravalle 2016 *(Continued)*

	Exclusion criteria: Women who did not perform defaecography and echodefaecography – *Additional data from authors received*
Patient characteristics and setting	**Nr of included patients:** 24 **Gender:** Female 100% **Age:** Mean age 57 (range 30 - 71) years **Symptoms:** ODS 100% (ODS Score medium 17) **Ethnicity:** *Additional data from authors received:* White **Co-morbidities:** – *Additional data from authors received:* Number of births median 2 (range 1 - 3), 0 previous anal surgery **Setting:** – *Additional data from authors received:* Secondary care, Single centre **Time period:** From May 2010 to May 2014 **Country study is conducted:** Argentina
Index tests	**Name index test:** Echodefaecography **Details of conducting index test:** – *Additional data from authors received:* Echodefaecography was performed with 2050 endoprobe (360 °) with 3 automatic scans (Flex Focus 1202, BK Medical, Denmark) **Imaging acquisition:** – *Additional data from authors received:* 3 automatic scans acquiring 3D volumes at rest, valsalva and evacuation. Operator with experience in anorectal ultrasound (10 - 20 years training). Rectal contrast was used and the participant was scanned in the left-lateral position **Imaging analysis:** – *Additional data from authors received:* 2 examiners, discrepancy meeting: yes, Blinded: yes **Threshold test positivity:** – *Additional data from authors received:* Rectocele: any; enterocele, below ischiococcigeal line; intussusception: protrusion of rectal wall layers during straining; anismus: Closure of anorecal junction angle during straining
Target condition and reference standard(s)	**Name index test 'EP':** Defaecography **Details of conducting evacuation proctography:** – *Additional data from authors received:* Philips X-ray machine. Defaecography was performed after inserting 150 ml of barium paste in the rectum **Imaging acquisition:** – *Additional data from authors received:* At rest, squeeze and evacuation, by experienced operator **Imaging analysis:** – *Additional data from authors received: 2* examiners, discrepancy meeting: yes, Blinded: yes **Threshold test positivity:** – *Additional data from authors received:* Rectocele: any; enterocele: below pubococcygeal line; intussusception: Invagination of the rectal wall during straining; anismus: closure of anorecal junction angle during straining
Flow and timing	**Enrolment and exclusions (+ reasons):** All 24 were included in 2 x 2 table **Nr analysed:** 24 **Time interval (+ interventions) between index test and reference standard:** *Additional data from authors received:* Interval < 30 days, no interventions between evaluations -

Miravalle 2016 *(Continued)*

Comparative
Notes

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		

Miravalle 2016 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Murad-Regadas 2008

Study characteristics	
Patient Sampling	**Patient selection:** Female patients with obstructed defaecation symptoms were prospectively enrolled in the study. The participants were submitted to a complete proctological examination, followed by defaecography (DF) and echodefaecography (EDF) performed by different examiners **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of the present study was to test echodefaecography (EDF), a novel 3D dynamic ultrasonography technique using ultrasound gel in the rectum to assess OD patients, and compare it to conventional defaecography (DF) **Inclusion criteria:** Female patients with obstructed defaecation symptoms **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 30 **Gender:** Female (100%) **Age:** Median age of 47.7 years (range 24 – 79 years) **Symptoms:** ODS 100%. Mean validated Wexner constipation score of 14 (range 7 – 25) (SD ± 4.66). **Ethnicity:** Unknown **Co-morbidities:** Unknown **Setting:** Tertiary, single centre **Time period:** March and November 2006

Murad-Regadas 2008 *(Continued)*

	Country study is conducted: Brazil
Index tests	**Name index test:** Echodefaecography **Details of conducting index test:** EDF was performed with a 3D ultrasound machine (Hawk, endoprobe model 2050, B-K Medical1, Herlev, Denmark) with a proximal-to-distal 6.0-cm automatic scan of 50 seconds, resulting in a 3D volume displayed as a cube and recorded and analysed in multiple planes. Participants were examined in the left lateral position after rectal enema. The endoprobe was inserted into the lower rectum and positioned 6 – 7 cm from the anal verge **Imaging acquisition:** Images were acquired by 3 automatic scans and analysed in axial and 3D midline longitudinal (ML) planes by an examiner blinded to the DF findings Scan 1: (at rest position without gel) was performed to visualise the anatomic integrity of the anal sphincters Scan 2: (at rest–straining–at rest without gel) evaluated the voluntary muscle relaxation during the evacuatory effort to identify anismus as demonstrated in a previous publication Scan 3: (at rest–straining–at rest) was performed by inserting ultrasound gel (120 – 180 ml) into the rectum. Participants remained quiet during the first 15 seconds, strained maximally for 20 seconds, then relaxed again. The scanning process was repeated up to 3 times, refilling the rectum with ultrasound gel whenever an image re-evaluation was required. In normal patients, the posterior vaginal wall displaced the lower rectum and the upper anal canal downwards and backwards, but the same straight horizontal position was maintained during the entire defaecation effort. **Imaging analysis:** Blinded **Threshold test positivity:** Rectocele: not-prespecified; Intussusception: presence; anismus: presence
Target condition and reference standard(s)	**Name index test 'EP':** Conventional defaecography (DF) **Details of conducting evacuation proctography:** DF was performed without prior bowel preparation. No opacification of the small bowel was performed. Opacification of the vagina with 50 mL barium paste and sufficient contrast filling of the rectum (300 mL barium paste) **Imaging acquisition:** The participant was asked to sit on a special commode, contract the pelvic floor musculature and empty the rectum as completely as possible **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: all stages (grade 1: < 2.0 cm; grade II: 2.0 – 4.0 cm; grade III: > 4.0 cm); intussusception: unknown; anismus: unknown; pelvic floor descent: unknown
Flow and timing	**Enrolment and exclusions (+ reasons):** All received both tests and included in 2 x 2 table **Nr analysed:** 30 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* About 1 week
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			

Murad-Regadas 2008 *(Continued)*

Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	No		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			

Murad-Regadas 2008 *(Continued)*

Are there concerns that the target condition as defined by the reference standard does not match the question?		Low concern
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Murad-Regadas 2011

Study characteristics	
Patient Sampling	**Patient selection:** Patients with obstructed defaecation disorder were prospectively enrolled in the study. The participants were given a complete proctological examination and underwent defaecography and dynamic 3-DAUS performed by different examiners for the evaluation of pelvic floor dysfunctions **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The purpose of the study was to describe a novel 3-dimensional dynamic anorectal ultrasonography technique (dynamic 3-DAUS) for assessment of perineal descent (PD) and establishment of normal range values, comparing it with defaecography **Inclusion criteria:** Patients with obstructed defaecation disorder (excessive straining, vaginal splinting and sensation of incomplete evacuation), despite increased intake of dietary fibre (up 30 g/day for 3 months) **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 29 **Gender:** Female (100%) **Age:** Mean age 43 years, range 23 – 74 **Symptoms:** ODS 100%. Mean validated Wexner constipation score of 10 (range 8– 18) **Ethnicity:** Unknown **Co-morbidities:** 15 participants had undergone at least 1 vaginal delivery **Setting:** Tertiary, single centre **Time period:** March 2008 and February 2009 **Country study is conducted:** Brazil
Index tests	**Name index test:** Dynamic 3-DAUS (Dynamic anorectal ultrasonography)

Murad-Regadas 2011 *(Continued)*

	Details of conducting index test: Dynamic 3-DAUS was performed with a 3-dimensional ultrasound device (Pro-Focus, endoprobe model 2052; B-K Medical, Herlev, Denmark) with proximal to distal 6.0-cm automatic scans. By moving 2 crystals on the extremity of the transducer, axial and longitudinal images were merged into a single cube image, recorded and analysed in multiple planes, as described in previous publications by Murad-Regadas et al. Following rectal enema, participants were examined in the left lateral position **Imaging acquisition:** Images were acquired by 4 automatic scans and analysed in the axial, sagittal and, if necessary, the oblique plane by an examiner blinded to defaecography findings. Scans 1, 3 and 4 used a slice width of 0.25 mm and lasted 55 seconds each. Scan 2 lasted 30 seconds with a slice width of 0.35 mm Scan 1: (rest) For verification of the anatomic integrity of the anal sphincters Scan 2: The transducer was positioned proximally to the PR (anorectal junction). The scan started with the participant at rest (3.0 seconds), followed by maximum straining with the transducer in fixed position. When the PR became visible distally, the scan was stopped. Perineal descent was quantified by measuring the distance between the position of the proximal border of the PR at rest and the point to which it had been displaced by maximum straining(PR descent). Instead of using the cut-off value for defaecography (> 3 cm), normal-range values were established for dynamic 3-DAUS by comparing with measurements from defaecography Scan 3: The transducer was positioned at 6.0 cm from the anal verge. The participant was requested to rest during the first 15 seconds, strain maximally for 20 seconds, then relax again, with the transducer following the movement. The purpose of the scan was to evaluate the movement of the PR and the external anal sphincter during straining, identifying normal relaxation, non-relaxation or paradoxical contraction (anismus). Scan 4: Following injection of 120–180 ml ultrasound gel into the rectal ampulla, the transducer was positioned at 7.0 cm from the anal verge. The scanning sequence was the same as in scan 3. The purpose of the scan was to visualize and quantify all anatomical structures and functional changes associated with voiding (rectocele, intussusception, Grade III sigmoidocele/enterocele). **Imaging analysis:** Blinded **Threshold test positivity:** Rectocele: any (0.2–0.6 cm for grade I, 0.7–1.3 cm for grade II, and >1.3 cm), Intussusception: any, Anismus: decrease of ARA during straining, Pelvic floor descent: not pre-defined.
Target condition and reference standard(s)	**Name index test:** Defaecography **Details of conducting evacuation proctography:** Defaecography was performed without opacification of the small bowel. Following rectal enema, the participant was placed in the left lateral position and approximately 25 ml of liquid iodine contrast was introduced into the vagina in order to demonstrate the effect of defaecation on the posterior vaginal wall. The rectum was filled with 200 ml barium paste **Imaging acquisition:** The participaent was asked to sit on a special commode, contract the pelvic floor musculature and empty the rectum as completely as possible. Measurements were made at rest, squeeze and during expulsion of the contrast **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele: any (Grade I < 2.0 cm; Grade II 2.0 – 4.0 cm; Grade III > 4.0 cm); enterocele: small bowel below ischiococcygeal line; Intussusception: any; anismus: muscles failed to relax or contracted during defaecation; pelvic floor descent: A difference of > 3 cm in the position of the anal canal between relaxation and straining
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants included underwent defaecography and dynamic 3-DAUS and all were included in the 2 x 2 table **Nr analysed:** 29 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* About 1 week

Murad-Regadas 2011 *(Continued)*

Comparative			
Notes			
Methodological quality			
Item	**Authors' judgement**	**Risk of bias**	**Applicability concerns**
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity prespecified?	No		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or inter-			Low concern

Murad-Regadas 2011 *(Continued)*

pretation differ from the review question?			
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Perniola 2008

Study characteristics	
Patient Sampling	**Patient selection:** 37 consecutive patients with obstructed defaecation were recruited from October 2005 to March 2007 and were scheduled to undergo defaecation proctography. In 34 women this involved an additional ultrasound examination; 3 participants authorised us to use previously acquired ultrasound data. In most cases the defaecation proctogram was carried out first. In 4 participants the order was reversed

Perniola 2008 *(Continued)*

	Study design: Cross-sectional test accuracy study, prospective (+ 3 cases retrospective) **Study objective:** The aim of this comparative study was to determine agreement between translabial ultrasound and defaecation proctography findings **Inclusion criteria:** Women with obstructed defaecation **Exclusion criteria:** Unknown
Patient characteristics and setting	**Nr of included patients:** 37 **Gender:** Female (100%) **Age:** Mean age 53 years (range 26 – 80) **Symptoms:** Constipation 26 (70%), straining at stool 31 (84%), vaginal digitation 15 (41%), sensation of incomplete emptying 30 (81%), Faecal incontinence 10 (27%), vaginal lump 10 (27%) **Ethnicity:** Unknown **Co-morbidities:** Median vaginal parity was 2 (range, 0 – 6). The mean age at first delivery was 24 (range, 17 – 39) years. 6 participants out of 37 were nulliparous. 10 women (27%) had had a previous hysterectomy, and 4 (11%) repair of a vaginal prolapse. 2 had previously undergone surgery for obstructed defaecation **Setting:** Tertiary care, single centre **Time period:** October 2005 to March 2007 **Country study is conducted:** Australia
Index tests	**Name index test 'miscellaneous':** Translabial Ultrasound **Details of conducting index test and image acquisition:** Translabial ultrasound was performed using a GE Kretz Voluson 730 Expert system (GE Medical, Sydney, Australia), after voiding, supine, at rest and on maximal Valsalva manoeuvre, as previously described. The procedure was noninvasive, as we did not use a contrast medium **Imaging analysis:** Volume data were archived and analysed at a later date by an operator blinded to all clinical data and defaecation proctography results **Threshold test positivity:** Rectocele: 10 mm depth; intussusception: presence/absence
Target condition and reference standard(s)	**Name index test 'EP':** Defaecation proctography **Details of conducting evacuation proctography:** Multiple fluoroscopic images were acquired using a Philips MD3 digital C-arm X-ray machine (Philips Healthcare, North Ryde, NSW, Australia). Thin barium or liquid polybar plus was instilled into the rectum in the first pass followed by a liquid polybar/starch mixture **Imaging acquisition:** Images were acquired at rest, during straining, defaecation and coughing, and the procedure was videotaped **Imaging analysis:** Measurements were obtained by different operators blinded to all clinical and imaging data; i.e. the person evaluating the ultrasound scan was unaware of the findings of the clinician reporting on the defaecation proctogram, and vice versa **Threshold test positivity:** Rectocele: any; intussusception: any
Flow and timing	**Enrolment and exclusions (+ reasons):** 37 consecutive patients with obstructed defaecation were recruited from October 2005 to March 2007. 6 women (16%) did not attend their defaecation proctography. 1 participat was found to be pregnant and 5 others cancelled repeatedly, leaving 31 cases for comparison. 1 of these had an incomplete proctogram owing to inability to defaecate, allow-

Perniola 2008 *(Continued)*

	ing only assessment for rectocele. All underwent 4-dimensional pelvic floor ultrasound examination **Nr analysed:** 31 **Time interval (+ interventions) between index test and reference standard:** The median interval between the 2 tests was 28 (range, 0 – 198) days.
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern

Perniola 2008 *(Continued)*

DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Pilkington 2012

Study characteristics	
Patient Sampling	**Patient selection:** All participants had been seen at a clinic in Poole or Dorchester as part of the pelvic floor service and had been referred for BaP as part of their National Health Service (NHS) management. This study invited 216 patients to participate. At the appointment for BaP, 71 participants were recruited and 42 of these completed the study by attending for MR proctography. The remaining 29 patients withdrew from the study. **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** The aim of this study was to compare BaP and MR proctography in the same individual to see if there were measurable differences between the 2 tests for clinically relevant findings **Inclusion criteria:** Referred for proctography as part of routine NHS management. Participant gives informed written consent. Participant is > 18 years old

Pilkington 2012 *(Continued)*

	Exclusion criteria: Patient incompetent to give informed consent. Claustrophobia or unable to tolerate MRI. Contraindications to MRI such as pacemaker, high body mass index. Patient unable to lie flat
Patient characteristics and setting	**Nr of included patients:** 42 **Gender:** 38 (90%) women and 4 (10%) men **Age:** The mean age of participants was 59 years with a range of 37 – 76 years **Symptoms:** Symptomatic pelvic floor disorders **Ethnicity:** Unknown **Co-morbidities:** Unknown **Setting:** Secondary care, single centre **Time period:** Between 8 May 2008 and 11 December 2009 **Country study is conducted:** United Kingdom
Index tests	**Name index test:** MR proctography **Details of conducting index test:** The technique for MR proctography was similar to BaP in that the participant had contrast (ultrasound gel) placed in the rectum. However, no contrast was placed in the vagina or small bowel. The MRI scanner had a 1 T magnet (Phillips Intera). The participant was positioned supine during scanning with a support for the feet so that the knees and hips were flexed **Imaging acquisition:** The MR sequence was recorded over a 40-second time period while the participant attempted rectal evacuation whilst lying in the scanning machine. 20 T2-weighted single midsagittal sections each 5 mm thick were taken at 2-second intervals to build up a dynamic sequence as the participant was bearing down and evacuating the rectum **Imaging analysis:** MR proctography was reported by a consultant radiologist with pelvic floor subspecialisation. At the time of reporting, each radiologist was blinded to the results of the other proctogramme **Threshold test positivity:** Rectocele: any; enterocele; any; intussusception: any; anismus: no rectal contrast evacuated or persistent puborectalis spasm
Target condition and reference standard(s)	**Name index test 'EP':** Barium proctography (BaP) **Details of conducting evacuation proctography:** During BaP, the rectum was filled with contrast (barium paste). The vagina and small bowel were opacified with contrast medium. The participant was seated on a radiolucent commode behind a screen **Imaging acquisition:** Fluoroscopic images were taken in the sagittal plane during rest, contraction and rectal evacuation **Imaging analysis:** BaP was reported by a consultant radiologist with pelvic floor subspecialisation. At the time of reporting, each radiologist was blinded to the results of the other proctogramme **Threshold test positivity:** Rectocele: any; enterocele; any; intussusception: any; anismus: no rectal contrast evacuated or persistent puborectalis spasm
Flow and timing	**Enrolment and exclusions (+ reasons):** All that had both BaP and MR proctography were included in 2 x 2 table. **Nr analysed:** 42

Pilkington 2012 *(Continued)*

	Time interval (+ interventions) between index test and reference standard: Interval average 84 days (range 22 to 284 days). No interventions
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		

Pilkington 2012 *(Continued)*

Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Poncelet 2017

Study characteristics	
Patient Sampling	**Patient selection:** 50 women with a mean age of 65.5 years (range:53 - 72 years) who underwent defaecography and MRI between December 2006 and August 2009 for clinical suspicion of posterior compartment dysfunction, were included in this retrospective study **Study design:** Retrospective cross-sectional test accuracy study **Study objective:** The goal of this study was to compare conventional X-ray defaecography and dynamic MR defaecography in the diagnosis of pelvic floor prolapse of the posterior compartment **Inclusion criteria:** Women who underwent defaecography and MRI between December 2006 and August 2009 for clinical suspicion of posterior compartment dysfunction **Exclusion criteria:** Not reported
Patient characteristics and setting	**Nr of included patients:** 50 **Gender:** female (100%) **Age:** 65.5 years**(range** 53 - 72 years) **Symptoms:** Clinical suspicion of posterior compartment dysfunction

Poncelet 2017 *(Continued)*

	Ethnicity: Not reported **Co-morbidities:** Not reported **Setting:** Not reported **Time period:** Between December 2006 and August 2009 **Country study is conducted:** France
Index tests	**Name index test:** MR defaecography **Details of conducting index test:** Before the examination, the vagina and rectum were filled with 50 ml and 200 - 250 ml of ultrasound gel respectively. All MR defaecography examinations were performed with a closed magnet MR imaging 1.5-T unit (Signa®, General Electric Healthcare Milwaukee, WI, USA, or Achieva®, Philips Best, The Netherlands) **Imaging acquisition:** Participants were imaged in supine position, with knees slightly bent to facilitate defaecation, with a body phased array coil. The study protocol included static T2-weighted fast spin echo (FSE) sequences in the sagittal, coronal and transverse planes (TR, 4920 ms; TE, 24 ms; FOV, 24 cm; section thickness, 4-mm; matrix size, 288 × 256;NEX, 4 excitations), followed by fast, dynamic single-slice T2-weighted gradient echo sequences acquired in the sagittal plan every second during 80 seconds during progressive straining (TR, 32 ms; TE, 1.1 ms; flip angle, 55 °; FOV, 30 cm;slice thickness, 10-mm; matrix size, 384 × 256; NEX, 2 excitations) until evacuation of the gel. Then, 3-multislice T2-weighted gradient echo sequences were acquired in the sagittal, coronal, and transverse plane during maximum straining of 20 seconds (TR, 20 ms; TE, 1.1 ms; flip angle,55 °; 12 contiguous slices; FOV, 30 cm; matrix size, 320 × 256;NEX, 2 excitations) **Imaging analysis:** MR defaecography examinations were analysed by an experienced radiologist **Threshold test positivity:** Rectocele > 25 mm; enterocele: any; intussusception: any; anismus: any
Target condition and reference standard(s)	**Name index test 'EP':** X-ray defaecography **Details of conducting evacuation proctography:** Participants were given the evening before or 1 hour before the examination a Normacol® enema (Normacol Lavement®,Norgine Pharma, Amsterdam, The Netherlands). Then, each participant swallowed 400 ml of a barium solution in order to opacify the small bowel. 120 millilitre of barium sulfate mixed with Smecta®(Ipsen Pharma,Boulogne-Billancourt, France) was injected into the rectum and 20 ml of Microtrast®(barym sulfate, Guerbet,Roissy-Charles de Gaulle, France) in the vagina. A foleycatheter was inserted to localise the anal canal and the anorectal junction **Imaging acquisition:** 2 lateral images were acquired in the standing position, the first 1 at rest, and the second during straining without evacuation. Cineloop mode images were recorded during defaecation with the participant in sitting position **Imaging analysis:** Conventional X-ray defaecography examinations were analysed by 1 experienced gastroenterologist **Threshold test positivity:** Rectocele > 25 mm; enterocele: upper part of the vagina; Intussusception: any; anismus: any
Flow and timing	**Enrolment and exclusions (+ reasons):** All included in 2 x 2 tables **Nr analysed:** 50 **Time interval (+ interventions) between index test and reference standard:** Not described
Comparative	
Notes	

Poncelet 2017 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Unclear		
Did the study avoid inappropriate exclusions?	Unclear		
Could the selection of patients have introduced bias?		Unclear risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Unclear		
Could the conduct or interpretation of the index test have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		

Poncelet 2017 *(Continued)*

Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Regadas 2011

Study characteristics	
Patient Sampling	**Patient selection:** Women presenting with obstructed defaecation symptoms at 6 centres for colorectal surgery (3 in Brazil, 1 in Texas, 1 in Florida, and 1 in Venezuela) were initially and prospectively evaluated with a clinical examination consisting of a full proctologic evaluation, followed by defaecography and echodefaecography performed by different examiners across the 6 centres **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** This study was designed to validate the effectiveness of echodefaecography compared with defaecography in the assessment of anorectal dysfunctions related to obstructed defaecation **Inclusion criteria:** Women with symptoms of obstructed defaecation **Exclusion criteria:** Patients with previous anorectal and vaginal surgery, faecal incontinence, or previous anorectal radiation or both were excluded
Patient characteristics and setting	**Nr of included patients:** 86 **Gender:** Female (100%) **Age:** The median age was 53.4 (range, 26 – 77) years. **Symptoms:** ODS (100%).The median validated Wexner constipation score was 13.4 (range, 6 – 23) **Ethnicity:** Unknown

Regadas 2011 *(Continued)*

	Co-morbidities: Among the participants, 16 (18.6%) were nulliparous, 40 (46.5%) had had vaginal deliveries, and 30 (34.9%) had undergone Cesarean section **Setting:** Tertiary, multicentre **Time period:** January 2009 through October 2009 **Country study is conducted:** Brazil, Venezuela, USA (Texas, Florida)
Index tests	**Name index test:** Echodefaecography **Details of conducting index test:** Participants were examined in the left lateral position after rectal enema. Echodefaecography was performed with a Pro Focus 3-dimensional ultrasound scanner (B-K Medical, Herlev, Denmark) using a 2050 endoprobe with 55-second proximal-to-distal 6.0-cm automatic scanning, a frequency range of 10 MHz to 16 MHz, and a focal distance of 2.8 cm to 6.4 cm **Imaging acquisition:** With the probe positioned in the rectum at 6.0 cm to 7.0 cm from the anal verge, 3 automatic scans (50-second duration each) were performed to identify the anatomic changes during straining (20-second interval) Scan 1: (at-rest position without gel) was performed to visualise the anatomic integrity of the anal sphincter musculature and to evaluate the position of the external anal sphincter and puborectalis muscles at rest. The angle formed between a line traced along the internal border of the external anal sphincter/puborectalis muscles, and a line traced perpendicular to the axis of the anal canal was measured, as previously reported. Scan 2: (at rest–straining–at rest without gel) evaluated voluntary muscle movement during the evacuatory effort to identify the presence of normal relaxation. The participant was asked to rest during the first 15 seconds, strain maximally for 20 seconds, and rest again during the remaining 15 seconds of the scan. The resulting positions of the external anal sphincter/puborectalis muscles (represented by the angle size) were compared between scans 1 and 2 Scan 3: 120 mL to 180 mL of ultrasound gel was inserted into the rectum and the rest–strain–rest sequence, identical to scan 2, was performed. In normal patients, the posterior vaginal wall displaces the lower rectum and upper anal canal inferiorly and posteriorly but maintains a straight horizontal position during defaecatory effort **Imaging analysis:** Images were analysed in the axial and sagittal planes by an examiner blinded to the defaecography findings **Threshold test positivity:** Rectocele: any (grade I (< 6.0 mm), grade II (6.0 – 13.0 mm), or grade III (> 13.0 mm)); enterocele: small bowel was positioned below the pubococcygeal line; anismus: the ARA decreased by a minimum of 1 degree during valsalva
Target condition and reference standard(s)	**Name index test 'EP':** Defaecography **Details of conducting evacuation proctography:** A Fleet enema was initially administered. After vaginal opacification with a mixture of iodine and ultrasound gel, 150 mL of barium paste was inserted into the rectum. The participant was seated parallel to the X-ray table for lateral visualisation of the anal canal and rectum **Imaging acquisition:** Images were taken both at rest and during straining. The coccyx, sacrum, head of the femur, posterior wall of the rectum, and anal canal were identified. A normal defaecogram at rest showed the rectum angled posteriorly and parallel to the presacral space. The participant was asked to contract the pelvic floor musculature and empty the rectum as completely as possible **Imaging analysis:** Unknown **Threshold test positivity:** Rectocele:**a**ny (grade I (< 2.0 cm), grade II (2.0 – 4.0 cm), or grade III (> 4.0 cm)); enterocele: small bowel below the ischiococcygeal line (Grade 3); intussusception: any; anismus: paradoxical contraction present

Regadas 2011 *(Continued)*

Flow and timing	**Enrolment and exclusions (+ reasons):** All participants included underwent defaecography and echodefaecography and were included in the 2 x 2 table **Nr analysed:** 86 **Time interval (+ interventions) between index test and reference standard:** *Additional information from authors:* About 1 week
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			

Regadas 2011 *(Continued)*

For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Unclear		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Unclear risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			High
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Ron 2012

Study characteristics			

Ron 2012 *(Continued)*

Patient Sampling	**Patient selection:** Consecutive patients that underwent both DTPU and EP for obstructed defaecation were reviewed **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** To assess the value of DTPU compared with EP in women with defaecation disorders **Inclusion criteria:** *Additional information from the authors:* Patients age > 18 with symptoms of obstructed defaecation **Exclusion criteria:** *Additional information from the authors:* Pregnancy
Patient characteristics and setting	**Nr of included patients:** 102 **Gender:** *Additional information from the authors:* 81 (79%) female, 21 (21%) male **Age:** Unknown **Symptoms:** Symptoms of obstructed defaecation **Ethnicity:** *Additional information from the authors:* White **Co-morbidities:** Unknown **Setting:** Secondary care, single centre **Time period:** January 2008 and November 2010 **Country study is conducted:** Israel
Index tests	**Name index test:** Transperineal Ultrasound **Details of conducting index test:** *Additional information from the authors:* Transperineal ultrasound was performed using a small convex probe (Hitachi, Hi Vision, model Ascendus) positioned in the perineal body region. The participant was scanned in left-lateral decubitus position after insertion of 120 ml ultrasound gel (Medi Pharm US gel) **Imaging acquisition:** *Additional information from the authors:* Imaging during the dynamic process of squeeze and evacuation. The tests were performed by 2 different operators, each unaware of the results of the other **Imaging analysis:** Unknown **Threshold test positivity:** Unknown
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography:** *Additional information from the authors:* A barium paste (E-Z-paste / E-Z HD intermixed with "thick and easy" to create a 200 ml paste suspension) was introduced to the rectum by Hi Vac syringe. The vaginal orifice was imaged by barium-soaked tampon introduced by each examiner **Imaging acquisition:** The tests were performed by two different operators, each unaware of the results of the other. *Additional information from the authors:* The actual dynamic imaging was performed while the participant is seated of a specially-designed commode. Resting, squeeze and evacuation process were recorded and anatomical markers were evaluated. We use a Philips mobile C-arm X-ray machine **Imaging analysis:** Unknown **Threshold test positivity:** Unknown

Ron 2012 *(Continued)*

Flow and timing	**Enrolment and exclusions (+ reasons):** All participants underwent DTPU and EP and were included in the 2 x 2 table **Nr analysed:** 102 **Time interval (+ interventions) between index test and reference standard:** *Additional information from the authors:* Up to a month
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Ron 2012 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Steensma 2010

Study characteristics	
Patient Sampling	**Patient selection:** All women with symptoms related to posterior-compartment prolapse referred to our tertiary pelvic floor unit were included in this prospective observational study **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To assess the level of agreement between EP and 3DTPUS in diagnosing posterior compartment prolapse in participants with related symptoms **Inclusion criteria:** Women with symptoms related to posterior compartment prolapse **Exclusion criteria:** None
Patient characteristics and setting	**Nr of included patients:** 75 **Gender:** Female (100%) **Age:** The median age was 59 years (range: 22 – 83 years)

Steensma 2010 *(Continued)*

	Symptoms: All with symptoms related to posterior compartment prolapse. These included pelvic discomfort 50 (67%), obstructed defaecation 36 (48%) and faecal incontinence 26 (35%) or a combination **Ethnicity:** Not described **Co-morbidities:** A previous hysterectomy had been carried out in 31 (41%) women and a previous pelvic organ prolapse repair in 37 (49%). Median parity was 2 (range: 0 – 10 vaginal deliveries; 4 nulliparous women) **Setting:** Tertiary, single centre **Time period:** September 2005 and July 2007 **Country study is conducted:** The Netherlands
Index tests	**Name index test:** Transperineal Ultrasound **Details of conducting index test:** 3D transperineal ultrasound was performed using a GE Kretz Voluson 730 expert system (GE Healthcare, Clinical Systems, Hoevelaken, the Netherlands), using an abdominal 4–8 MHz transducer. Participants were examined after voiding and in the supine position **Imaging acquisition:** 2D cineloop volumes (3D) were obtained at rest, during levator contraction and during maximal Valsalva manoeuvre as previously described by Dietz 2005a. **Imaging analysis:** Offline evaluation of the cineloop volumes was performed by 1 gynaecologist (ABS), blinded against all clinical data and the results of EP, using 4D VIEW software (GE Healthcare) **Threshold test positivity:** Rectocele ≥ 10 mm depth; any enterocele; any intussusception
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography **Details of conducting evacuation proctography:** Evacuation proctography was performed using a standardised technique with opacification of the rectosigmoid, small bowel and vagina using liquid barium contrast **Imaging acquisition:** Imaging was acquired at rest, during pelvic floor contraction and during straining, and a video recording was obtained during evacuation of contrast **Imaging analysis:** All video files were analysed by 1 colorectal surgeon (WRS), blinded against all clinical data and the results of 3DTPUS **Threshold test positivity:** any rectocele depth; any enterocele; any intussusception
Flow and timing	**Enrolment and exclusions (+ reasons):** All enrolled received both imaging and no exclusions from 2 x 2 table **Nr analysed:** 75 **Time interval (+ interventions) between index test and reference standard:** EP and dynamic 3DTPUS were done with a maximum interval of 6 months
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns

Steensma 2010 *(Continued)*

DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the refer-			Low concern

Steensma 2010 *(Continued)*

ence standard does not match the question?		
DOMAIN 4: Flow and Timing		
Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Van Gruting 2017

Study characteristics	
Patient Sampling	**Patient selection:** In this prospective cohort study, between January 2014 and January 2015, consecutive women with symptoms of obstructed defaecation syndrome were recruited from tertiary urogynaecology or colorectal clinics in Croydon University Hospital. Evacuation proctography and MRI proctogram were requested as part of the hospital's protocol and additional transperineal and endovaginal ultrasonography were performed as part of this study **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** To establish the diagnostic test accuracy of evacuation proctography, MRI, and transperineal and endovaginal ultrasonography for the detection of posterior pelvic floor disorders in women with obstructed defaecation syndrome **Inclusion criteria:** Women with obstructed defaecation syndrome referred for evacuation proctography **Exclusion criteria:** Age younger than 18 years, inability to understand English, and lacking mental capacity
Patient characteristics and setting	**Nr of included patients:** 131 **Gender:** Female 100% **Age:** Mean age was 54 years (range 25 - 90) **Symptoms:** ODS 100%; 114 women (87%) had the feeling of incomplete emptying on a weekly basis, 62 (47%) had to digitally assist evacuation either vaginally or anally at least weekly, and 62 (47%) had to strain excessively at least 50% of the time **Ethnicity:** 77% white, 8% Asian and 15% black **Co-morbidities:** Body mass index 27 (64.9 SD) and parity 2.2 (61.3 SD). Previous pelvic organ prolapse surgery had been performed in 24 (18%) women, 39 (30%) had a hysterectomy, and 12 (9%) had previous surgery for obstructed defaecation syndrome (6 stapled transanal rectal resection, 4 rectopexy, 2 with both) **Setting:** Tertiary, single centre **Time period:** Between January 2014 and January 2015 **Country study is conducted:** United Kingdom

Van Gruting 2017 *(Continued)*

Index tests	**Name index test 1:** MR-proctography **Details of conducting index test:** MR-proctography was performed with a closed MRI scanner with a 1.5 T magnet (Siemens Avanto) by specifically trained radiographers. The rectum was filled with 120 ml of contrast (ultrasound gel) . The participant was scanned in supine position with the knees and hips flexed to facilitate evacuation of contrast **Imaging acquisition:** T2-weighted fast acquisition images were obtained simultaneously in the midsagittal and coronal planes to evaluate pelvic organ descent and pelvic floor muscles motion while the participant was instructed through headphones to perform the rest–squeeze–relaxation– strain–evacuation manoeuvre and to empty the rectum as completely as possible **Imaging analysis:** Offline analysis of images was performed by 2 observers blinded to clinical and other imaging findings. In case of discrepancies, final diagnosis was made by a third observer (a radiologist with more than 30 years' experience in pelvic floor imaging) **Threshold test positivity:** Rectocele > 2 cm depth; enterocele: small bowel below PCL; intussusception: full thickness circumferential invagination; anismus: paradoxical pelvic floor contraction; pelvic floor descent: ARJ > 30 mm below the PCL at valsalva **Name index test 2:** Transperineal ultrasonography **Details of conducting index test:** Transperineal ultrasonography was performed by an experienced ultrasonographer using Profocus ultrasound scanner, with the participant in a supine position with hips and knees semiflexed. No vaginal or rectal contrast was used. A convex transducer (Type 8802, 3.5-6.0 MHz, focal range 10-135 mm, BK Medical, Denmark) was gently placed on the perineum in a vertical position **Imaging acquisition:** Images were acquired at rest, squeeze and maximum Valsalva. 3 Valsalva manoeuvres were recorded as a cineloop and the best cineloop was used for analysis **Imaging analysis:** Offline analysis of images was performed by 2 observers blinded to clinical and other imaging findings. In case of discrepancies, final diagnosis was made by a third observer (a urogynaecologist with more than 10 years of experience in pelvic floor ultrasonography) **Threshold test positivity:** Rectocele: > 10 mm depth; enterocele: small bowel visible below the posterior inferior margin of symphysis pubis; intussusception: full-thickness circumferential invagination; anismus: paradoxical pelvic floor contraction; pelvic floor descent: ARJ below posterior inferior margin of symphysis pubis **Name index test 3:** Endovaginal ultrasonography **Details of conducting index test:** Endovaginal ultrasonography was performed by an experienced ultrasonographer using of Profocus ultrasound scanner with the participant in a supine position with hips and knees semiflexed. No vaginal or rectal contrast was used. A high-resolution linear array transducer (Type 8838, 6–12 MHz, focal range 3 - 60 mm, contact surface 65 x 5.5 cm, BK Medical, Denmark) was placed in the vagina **Imaging acquisition:** Images were acquired at rest, squeeze, and maximum Valsalva. Three Valsalva manoeuvers were recorded as a cineloop and the best cineloop was used for analysis. **Imaging analysis:** Offline analysis of images was performed by two observers blinded to clinical and other imaging findings. In case of discrepancies, final diagnosis was made by a third observer [an urogynecologist with more than 10 years of experience in pelvic floor ultrasonography]. **Threshold test positivity:** Rectocele: > 10 mm depth, Enterocele: small bowel below visible in region of the rectovaginal septum, Intussusception: full thickness circumferential invagination, Anismus: paradoxical pelvic floor contraction, Pelvic floor descent: > 25 mm difference between position of ARJ at rest and valsalva
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation Proctogram **Details of conducting evacuation proctography:** Evacuation proctography was performed by an experienced radiologist with a special interest in pelvic floor. The small bowel was opacified with oral diluted barium 1 hour before the procedure and the rectum was prepared with glycerin suppositories. The rectum was filled with 120 mL barium paste (barium sulphate mixed with potato powder). The patient was sitting on a radiolucent commode with a metal ruler placed adjacent to the patient to calibrate the images for analysis.

Van Gruting 2017 *(Continued)*

	Imaging acquisition: Images were recorded in the sagittal plane at rest, during contraction, straining, and evacuation of contrast **Imaging analysis:** Offline analysis of images was performed by 2 observers blinded to clinical and other imaging findings. In case of discrepancies, final diagnosis was made by a third observer (a radiologist with more than 30 years' experience in pelvic floor imaging) **Threshold test positivity:** Rectocele > 2 cm depth; enterocele: small bowel below PCL; intussusception: full-thickness circumferential invagination; anismus: paradoxical pelvic floor contraction; Pelvic floor descent: ARJ > 30 mm below the PCL at valsalva
Flow and timing	**Enrolment and exclusions (+ reasons):** All women underwent evacuation proctography and endovaginal and transperineal ultrasonography. In 4 women, MRI was contraindicated and 5 women had no MRI because they declined for other reasons **Nr analysed:** MRI 122, EVUS 131, TPUS 131 **Time interval (+ interventions) between index test and reference standard:** Pelvic floor ultrasonography, consisting of transperineal and endovaginal ultrasonography, was performed at the same time. The time difference (median) between evacuation proctography and MRI was 11.5 days (range 0 – 92 days), between evacuation proctography and ultrasonography 3.0 days (range 0 – 58 days), and between MRI and ultrasonography 8.5 days (range 0 – 89 days)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			

Van Gruting 2017 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			

Van Gruting 2017 *(Continued)*

Was there an appropriate interval between index test and reference standard?	Yes	
Did all patients receive the same reference standard?	Yes	
Were all patients included in the analysis?	Yes	
Could the patient flow have introduced bias?		Low risk

Van Iersel 2017

Study characteristics	
Patient Sampling	**Patient selection:** All consecutive patients of 1 gastrointestinal surgeon (ECJC) with symptoms of pelvic floor dysfunction of the posterior compartment requiring radiological assessment between June 2010 and June 2011, prospectively underwent D-MRI and CD **Study design:** Prospective cross-sectional test accuracy study **Study objective:** The aim of this study was to compare D-MRI with dynamic CD as the reference standard with rectal evacuation assessed with the use of radiological contrast in patients with symptoms of prolapse of the posterior compartment of the pelvic floor **Inclusion criteria:** Patients of 1 gastrointestinal surgeon (ECJC) with symptoms of pelvic floor dysfunction of the posterior compartment requiring radiological assessment **Exclusion criteria:** Not reported
Patient characteristics and setting	**Nr of included patients:** 45 **Gender:** 39 female / 6 male **Age:** 64.3 (range 38 – 85) **Symptoms:** Faecal incontinence 18, obstructed defaecation 18, constipation 4, change defaecation < 3 months 5, faecal urgency 7 **Ethnicity:** Not reported **Co-morbidities:** Parity 2.4 (range 0– 5). Previous surgery: rectopexy 5, hysterectomy 19 (46.3), cystopexy 7, colporrhaphy, anterior 11, colporrhaphy, posterior 10, sphincter operation 1, stapled haemorrhoidectomy 3, other abdominal surgery 15 **Setting:** Single centre, secondary **Time period:** Between June 2010 and June 2011 **Country study is conducted:** The Netherlands
Index tests	**Name index test:** Dynamic MR defaecography (D-MRI).

Van Iersel 2017 *(Continued)*

	Details of conducting index test: All D-MRI imaging studies were performed on a 1.5-T closed magnet (Intera rel.2.6.3, Philips, Best, The Netherlands). All participants were imaged supine with a body-phased-array receiver coil (Torso-XL). The participant was asked to remain on a low-fibre diet 24 hours before the examination. To ensure adequate bladder-filling, the participant was asked to avoid micturition 2 hours before the examination. The vagina and rectum were filled with 50 and 200 ml, respectively, of warm ultrasonographic gel. **Imaging acquisition:** After an initial localiser in 3 different planes, the study protocol included a turbo spin echo (TSE) T2-weighted (T2W) axial sequence (voxel size 1 9 1.25 mm; 53 images; thickness 4 mm; repetition time (TR)/echo time (TE), 6,430/114; flip angle 90 °; turbo factor 15; scan time 3.10 min), a TSE T2W sagittal sequence (voxel size 1.0 9 1.2 mm; 35 images; thickness 4 mm; TR/TE, 846/11; flip angle 90 °; turbo factor 15; scan time 3.04 min) and a functional dynamic sequence with a balanced fast-field echo (FFE) T2W sequence sagittal during squeezing, pushing, evacuation and after evacuation (voxel size 1.8 9 1.4 mm; 60 images in total; 1.5 seconds per image; thickness 8 mm; TR/TE, 3.75/1.6; flip angle 45 °; scan time 1.32 minutes) through the midline. No micturition/voiding was pursued and did not occur during this series. The dynamic images of this last sequence were presented in cinematic form **Imaging analysis:** 2 radiologists (BGFH, IS) independently reviewed the D-MRI images **Threshold test positivity:** Rectocele > 20 mm; enterocele: below PCL; intussusception: full-thickness circumferential; perineal descent > 30 mm below PCL during straining
Target condition and reference standard(s)	**Name index test 'EP':** Dynamic conventional (entero-colpo) defaecography (CD) **Details of conducting evacuation proctography:** For small bowel contrast, 65 ml of thick barium paste (barium sulphate, E-Z-HD) mixed with water (515% wt/vol) and 5 ml microlax (sodium laureth sulphate/sodium citrate/sorbitol) were administered to each particiant by mouth 2 hours before the examination. The participant was asked to pass urine before the examination in order to avoid pelvic crowding. No bladder contrast was used. The distal sigmoid colon was opacified with 300 ml of barium paste (barium sulphate, Liquid Polibar) and water (35% wt/vol) instilled through rectum with a colon cannula. The rectum was opacified using 150 ml barium paste mixed with Metamucil to create a consistency similar to stool. The anal canal was also demonstrated by contrast during removal of the syringe used to inject the contrast and the vagina was opacified with 10 ml of barium paste. The participant was seated on a radiolucent commode with the fluoroscopic table vertically upright **Imaging acquisition:** A lateral radiograph was taken with the participant at rest. Cineradiography (2 images per second) was performed at rest and during puborectalis contraction, a Valsalva manoeuvre, squeezing, evacuation and after evacuation. For measurements of midline structures corrected for magnification, a radiopaque chain of beads 4.4 mm from each other was attached to the patient's anal cleft **Imaging analysis:** The CD examinations were independently reported by 2 different radiologists **Threshold test positivity:** Rectocele > 20 mm; enterocele: below PCL; intussusception: full-thickness circumferential; perineal descent > 30 mm below PCL during straining
Flow and timing	**Enrolment and exclusions (+ reasons):** Women without rectal evacuation of contrast on D-MRI or CD were excluded. Two women were excluded from further analysis because no rectal evacuation was achieved on D-MRI with no faecal obstruction in either. Two further women were excluded because of extensive anal sphincter damage with an inability to retain the rectal contrast in one and the presence of a pessary during D MRI in the other. The degree of ARJ descent could not be measured in 8 participants in the CD series because of an inability to draw the PCL or a difficulty in calculating accurate magnification. For this latter reason it was not possible to measure the depth of the rectocele in 5 of these participants **Nr analysed:** 41 **Time interval (+ interventions) between index test and reference standard:** Not described
Comparative	Not applicable
Notes	

Van Iersel 2017 *(Continued)*

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			

Van Iersel 2017 *(Continued)*

Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Unclear		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Unclear risk	

Vanbeckevoort 1999

Study characteristics	
Patient Sampling	**Patient selection:** 35 women with clinical evidence of pelvic floor descent were included in the study **Study design:** Cross sectional-test accuracy study, prospective **Study objective:** The purpose of this study was to compare fast dynamic magnetic resonance imaging (MRI) with colpo-cysto-defaecography (CCD) in the evaluation of pelvic floor descent in women

Vanbeckevoort 1999 *(Continued)*

	Inclusion criteria: Women with clinical evidence of pelvic floor descent **Exclusion criteria:** Not able to strain adequately
Patient characteristics and setting	**Nr of included patients:** 35 **Gender:** Female 100% **Age:** mean age 65.4 years; range 44 – 83 years **Symptoms:** Clinical evidence of pelvic floor descent **Ethnicity:** Unknown **Co-morbidities:** 9 participants had had a hysterectomy **Setting:** Tertiary care, single centre **Time period:** Unknown **Country study is conducted:** Belgium
Index tests	**Name index test:** Dynamic magnetic resonance imaging **Details of conducting index test:** All MR imaging studies were performed on a 1.5 T system (Magnetom Vision, Siemens Medical Systems, Erlangen, Germany) with a gradient switching capability of 25 mT/m in a rise time of 600 msec. All participants were imaged with a body-phased-array receiver coil. The rectum was filled with 100 mL of aqueous sonographic gel. No opacification of the bladder, the vagina, or small bowel was used **Imaging acquisition:** After an initial localiser in 3 different planes, half-Fourier single-shot turbo spin-echo (HASTE) images were obtained in the sagittal plane during pelvic floor relaxation and during maximal pelvic strain. The HASTE sequence is a T2-weighted acquisition in which all radiofrequency (RF) refocused echoes are obtained after a single excitation. The following parameters were used: time interval between subsequent echoes 4.2 msec; effective TE 60 msec; flip angle 160 °; number of excitations 1; matrix 160 3 256. Only half of the k-space was measured (echo train length 88); the k-space was then expanded with the half-Fourier method to 160 lines. Slice thickness was 5 mm, and 20 slices were obtained in 1 acquisition with a distance between the measured slices of 5 mm and during quiet breathing. The field of view (FOV) was 300 – 320 mm (with rectangular FOV 6/8 if possible) and receiver bandwidth was 650 Hz/pixel **Imaging analysis:** All personal information was removed from the radiological images. The images were then independently assessed by 2 experienced observers (DVB and LVH). In case of disagreement the final diagnosis was made by consensus **Threshold test positivity:** Rectocele: > 3 cm depth; enterocele: below PCL; pelvic floor descent: ARJ > 2.5 cm below PCL
Target condition and reference standard(s)	**Name index test 'EP':** Colpo-cysto-defaecography (CCD) **Details of conducting evacuation proctography:** A dynamic CCD was performed with the participant seated on a stool-chair. Opacification included the bladder (Telebrix), the vagina (Hytrast), and the rectum (barium). The small bowel was opacified by a barium meal 90 minutes prior to the CCD **Imaging acquisition:** Lateral images (100 mm camera) were obtained with conventional X-ray equipment (Diagnost 75, Philips-Fluorospot Siemens) at rest, during maximal pelvic strain, and during voiding and defaecation. The process of voiding and defaecation was recorded on videotape **Imaging analysis:** All personal information was removed from the radiological images. The images were then independently assessed by 2 experienced observers (DVB and LVH). In case of disagreement the final diagnosis was made by consensus

Vanbeckevoort 1999 *(Continued)*

	Threshold test positivity: Rectocele: > 3 cm depth; enterocele: below PCL; pelvic floor descent: ARJ > 2.5 cm below PCL
Flow and timing	**Enrolment and exclusions (+ reasons):** All enrolled participants were included in the 2 x 2 table **Nr analysed:** 35 **Time interval (+ interventions) between index test and reference standard:** On the same day
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or in-			Low concern

Vanbeckevoort 1999 *(Continued)*

terpretation differ from the review question?			
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Vitton 2011

Study characteristics	
Patient Sampling	**Patient selection:** Women with a history of dyschezia undergoing diagnostic evaluation at a regional referral centre in Marseille, France, between January 2009 and June 2010 were eligible for the study. Dyschezia was defined according to Rome III criteria by excessive straining, lumpy or hard stools, sensation of incomplete evacuation of stools, sensation of anorectal obstruction, manual disimpaction of stool, or vaginal manoeuvres to assist defaecation. All participants had constipation severe enough to be referred to a gastroenterologist in our tertiary centre for pelvic floor disorders **Study design:** Cross-sectional test accuracy study, prospective **Study objective:** This study aimed to compare the accuracy of dynamic anorectal endosonography and dynamic MRI defaecography with conventional defaecography as the criterion standard in the diagnosis of pelvic floor disorders

Vitton 2011 *(Continued)*

	Inclusion criteria: The inclusion criteria were female gender, age 18 years or older, dyschezia for at least 6 months, and willingness to participate in the study **Exclusion criteria:** The exclusion criteria were organic pathology of the colon or rectum detected by clinical examination or colonoscopy, pregnancy, anal incontinence, refusal to undergo 3 evaluations, previous surgery for pelvic floor disorders, contraindications to performance of MRI or DAE (for example, anal stenosis)
Patient characteristics and setting	**Nr of included patients:** 56 **Gender:** Female (100%) **Age:** The mean age was 50.7 (SD, 12.5; range, 25 – 80) years. **Symptoms:** Dyschezia for > 6 months (100%) **Ethnicity:** Unknown **Co-morbidities:** Hysterectomy in 6 (10.7%) **Setting:** Tertiary care, single centre **Time period:** January 2009 and June 2010 **Country study is conducted:** France
Index tests	**Name index test 1:** Dynamic MRI defaecography **Details of conducting index test:** MRI was performed with the participant in the supine position in a manner similar to that described by Kelvin 2000 with a 1.5-T superconductive unit and a circularly polarised (quadrature) body coil (Intera; Philips Medical Systems, Best, The Netherlands). The participants were asked to empty their bladder on arrival at the department. Before the beginning of the examination, the participants were instructed as to the voluntary manoeuvres to be performed during imaging. Manoeuvres consisted of progressive straining and contraction of the pelvic floor muscles (squeezing) followed by relaxation and rectal evacuation. Participants were given an explanation of the importance of rectal evacuation, emphasising that evacuation was essential to obtain complete information about the degree of prolapse. Waterproof padding was placed beneath the buttocks and thighs to limit participant embarrassment and to protect the table of the MRI unit. The rectum was opacified with 100 mL of sonographic transmission gel (Aquasonic 100; Parker Laboratories, Fairfield, NJ) introduced through a 26-French catheter **Imaging acquisition:** The participants were asked to perform the rest-squeeze-relax-strain-evacuate manoeuvre. During this process, a dynamic series of images was obtained in the midsagittal plane using true fast imaging in a steady-state free precession sequence (TR/TE, 6.32/3.00; flip angle, 70 °; matrix size, 192 256; field of view, 250 – 330 mm; 1 image every 1.2 seconds). The rest-squeeze- relax-strain-evacuate manoeuvre and the imaging were repeated so that imaging during complete rectal evacuation could be obtained **Imaging analysis:** All MRI assessments were done by the same experienced radiologist. Experienced senior operators without knowledge of the previous findings performed all measurements, and all measurements were recorded under blinded conditions on separate sheets **Threshold test positivity:** Rectocele: > 2 cm; enterocele: below PCL; intussusception: mucosal or full-thickness; perineal descent: ARJ below PCL **Name index test 2:** Dynamic anal endosonography **Details of conducting index test:** DAE was performed with the participant in the left lateral decubitus position. A rigid biplane transrectal probe with a frequency of 7 MHz was used (model EUP-U533; Hitachi Medical Systems, Tokyo, Japan). The tip of the probe was covered with a water-filled balloon to maintain the acoustic window for the ultrasound waves. The procedure was preceded by the filling of the rectum with 50 mLof water before the defaecation effort because we had demonstrated improved results with this procedure in a previous study **Imaging acquisition:** By slowly and manually rotating the linear probe through 360 °, we could identify the various layers constituting the anal wall (mucosa, internal anal sphincter, and external anal sphincter), the layer forming the rectal wall, and the perirectal tissues (puborectalis muscle, bladder, and vagina, or

Vitton 2011 *(Continued)*

	prostate). After the initial examination, the participants were asked to make a defaecation effort with the probe left in the same position **Imaging analysis:** All DAE measurements were done by the same operator. Experienced senior operators without knowledge of the previous findings performed all measurements, and all measurements were recorded under blinded conditions on separate sheets **Threshold test positivity:** Rectocele > 2 cm depth; enterocele: grade III; perineal descent: > 2 cm descent of puborectalis muscle on valsalva
Target condition and reference standard(s)	**Name index test 'EP':** Conventional defaecography **Details of conducting evacuation proctography:** Conventional defaecography was performed using a simplified method described by Mahieu 1984. The small bowel and the vagina were also opacified. After sufficient contrast filling of the rectum (300 mL), the participants were asked to sit on a special commode, contract the pelvic floor musculature, and then to empty the rectum as completely as possible **Imaging acquisition:** Fluoroscopic images were recorded during several such manoeuvres to assess and measure the descent of the pelvic floor and to diagnose any rectocele, enterocele, or rectal intussusception **Imaging analysis:** All these assessments were done by the same experienced radiologist, who did not perform the dynamic MRI defaecography. Experienced senior operators without knowledge of the previous findings performed all measurements, and all measurements were recorded under blinded conditions on separate sheets **Threshold test positivity:** Rectocele: > 2 cm; enterocele: below PCL; intussusception: full-thickness; perineal descent: ARJ > 3 cm below the PCL on straining
Flow and timing	**Enrolment and exclusions (+ reasons):** All participants included in the 2 x 2 table **Nr analysed:** 56 **Time interval (+ interventions) between index test and reference standard:** The 3 procedures were performed in random order within the same month
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			

Vitton 2011 *(Continued)*

Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	

Vitton 2011 *(Continued)*

If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Weemhoff 2013

Study characteristics	
Patient Sampling	**Patient selection:** Women with complaints of faecal incontinence or obstructed defaecation visiting the tertiary-care colorectal pelvic floor unit and who were scheduled to undergo a diagnostic endoanal ultrasonography and evacuation proctography were consecutively asked to join the study. The participants included underwent endoanal ultrasonography, transperineal ultrasonography, and evacuation proctography **Study design:** Prospective observational cross-sectional study **Study objective:** To determine the level of agreement between transperineal ultrasound and evacuation proctography **Inclusion criteria:** Women with complaints of faecal incontinence or obstructed defaecation visiting tertiary care **Exclusion criteria:** Age < 18 years, legally incapable, and persons who were not able to understand the information given
Patient characteristics and setting	**Nr of included patients:** 50 women were included in the study **Gender:** Female 100% **Age:** The mean age was 59 years (range 28 – 95).

Weemhoff 2013 *(Continued)*

	Symptoms: 82% of women had faecal incontinence, and 16% had complaints of obstructed defaecation **Ethnicity:** Not described **Co-morbidities:** Not described **Setting:** This was performed at the Maastricht University Medical Centre **Time period:** Between April 2007 and February 2008 **Country study is conducted:** The Netherlands
Index tests	**Name index test:** Transperineal ultrasound **Details of conducting index test:** Transperineal ultrasound was performed using a GE Kretz Voluson 730 expert system. For transperineal ultrasonography a 4- to 8-mHz transabdominal curved 2D transducer was used. No contrast medium was used. The transducer was placed against the perineum in the midsagittal plane with a maximum angle of 70 °. The ultrasound examinations were performed with the participant in the supine position with slightly flexed legs. Participants were requested to empty their bladders prior to the examination **Imaging acquisition:** Imaging was acquired at rest, during contraction, and during straining. At least 3 valsalva manoeuvres were recorded as a cineloop, and the best of these manoeuvres was used for evaluation **Imaging analysis:** The observers were blinded to other results (symptoms, physical examination, and other imaging studies). The datasets were anonymised. The ultrasound data were assessed independently by 2 experienced urogynaecological ultrasonographers (MW and KK). After establishing the interobserver agreement between the ultrasonographers, a consensus meeting was held on the cases the assessors disagreed about **Threshold test positivity:** Rectocele: > 2 cm depth; enterocele: any; intussusception: any
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctogram **Details of conducting evacuation proctography:** Evacuation proctography was performed using a standardised technique with opacification of the rectosigmoid. Orally, liquid barium contrast was given to make the small bowel visible. The vagina was filled with contrast to visualise the vagina **Imaging acquisition:** Imaging was performed at rest, during contraction, and during straining and evacuation of the contrast **Imaging analysis:** Evaluation of the recorded videos and photos was performed independently by 2 experienced observers. Disagreement between assessors was resolved at a consensus meeting **Threshold test positivity:** Rectocele > 2 cm; enterocele upper half of vagina; intussusception; any
Flow and timing	**Enrolment and exclusions from analysis:** 50 women were included in the study and all were included in the analysis **Reason for exclusions:** N/A **Nr analysed:** 50 **Time interval (+ interventions) between index test and reference standard:** Evacuation proctogram and transperineal ultrasound were performed on the same day
Comparative	
Notes	
Methodological quality	

Weemhoff 2013 *(Continued)*

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			High
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its inter-		Low risk	

Weemhoff 2013 *(Continued)*

pretation have introduced bias?			
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		
Were all patients included in the analysis?	Yes		
Could the patient flow have introduced bias?		Low risk	

Zafar 2012

Study characteristics	
Patient Sampling	**Patient selection:** Data were collected retrospectively by reviewing clinical letters, anorectal physiology reports, and radiology reports for patients with OD who underwent both MRD and EP between 2008 and 2011. There were 118 MRDs and 102 EPs performed at our institution during the study period. 16 participants underwent both diagnostic studies **Study design:** Cross-sectional test accuracy study, retrospective **Study objective:** The aim of this study is to compare supine magnetic resonance defaecography and evacuation proctography for the evaluation of the posterior pelvic compartment **Inclusion criteria:** Patients who underwent both diagnostic studies **Exclusion criteria:** Patients who did not undergo both diagnostic studies
Patient characteristics and setting	**Nr of included patients:** 16 **Gender:** Female: 13 (81%), Male: 3 (19%) **Age:** Mean age 39 years **Symptoms:** Common presenting symptoms were sensation of incomplete evacuation (93%), digitation (43%), faecal incontinence (31%), urgency (18%), and prolapse (18%) **Ethnicity:** Not described **Co-morbidities:** Not described **Setting:** Secondary care, single centre

Zafar 2012 *(Continued)*

	Time period: Between 2008 and 2011 **Country study is conducted:** United Kingdom
Index tests	**Name index test:** Magnetic resonance defaecography (MRD) **Details of conducting index test:** MRD examinations were performed on a 1.5 Tesla closed magnet Siemens Symphony scanner. The participant lies supine on a waterproof mat in the MRI scanner, with knees slightly flexed; legs apart and a pillow underneath. A flexible transmit/receive radiofrequency Siemens 6 channel multiphase coil is wrapped around the pelvis. Patient evacuates pre-instilled rectal contrast (ultrasound gel) on the MR table **Imaging acquisition:** – *additional information from authors:* During evacuation of the contrast **Imaging analysis:** – *additional information from authors:* 1 examiner, analysis not blinded (retrospective study) **Threshold test positivity:** *additional information from authors:* Rectocele: any; intussusception: any; anismus: present/absent
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation Proctography (EP) **Details of conducting evacuation proctography:** During EP the participants were seated on a commode, feet placed on the footrest of an upright-positioned examination table in front of a fluoroscopic unit. Participant evacuates pre-installed rectal contrast in a sitting position **Imaging acquisition:** *additional information from authors:* During evacuation of the contrast **Imaging analysis:** *additional information from authors:* 1 examiner, analysis not blinded (retrospective study) **Threshold test positivity:** *additional information from authors:* Rectocele: any; intussusception: any; anismus: present/absent
Flow and timing	**Enrolment and exclusions (+ reasons):** Of the included participants all 16 had both imaging techniques and were included in the 2 x 2 table **Nr analysed:** 16 **Time interval (+ interventions) between index test and reference standard:** The median interval between studies was 4.5 months (IQR: 2.25 to 11.25)
Comparative	
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	No		
Did the study avoid inappropriate exclusions?	No		
Could the selection of patients have introduced bias?		High risk	

Zafar 2012 *(Continued)*

Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and setting do not match the review question?			Low concern
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	No		
Could the conduct or interpretation of the index test have introduced bias?		High risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	No		
Could the reference standard, its conduct, or its interpretation have introduced bias?		High risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	No		
Did all patients receive the same reference standard?	Yes		

Zafar 2012 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	High risk

Zafar 2017

Study characteristics	
Patient Sampling	**Patient selection:** Patients with ODS were recruited from pelvic floor clinics across a single National Health Service Trust between the years 2012 and 2015. Patients who already had or were planned to have EP were invited to participate in the study and have an additional scan (MRD) after informed consent **Study design:** Prospective, cross-sectional test accuracy study **Study objective:** The aim of this prospective study was to compare the findings and acceptability of MRD and EP in the same cohort of patients. **Inclusion criteria:** Patients with symptoms of obstructive defaecation **Exclusion criteria:** Age < 18 or > 90, previous operations for obstructive defaecation, colorectal cancer, mentally incapacitated, do not understand English, for whom magnetic resonance imaging is contraindicated (pacemaker, aneurysmal clips), positive pregnancy test
Patient characteristics and setting	**Nr of included patients:** 55 **Gender:** 53 female, 2 male *(DTA data received on women only)* **Age:** 59 (interquartile range 50 - 65) **Symptoms:** ODS **Ethnicity:** Not reported **Co-morbidities:** Not reported **Setting:** Secondary, single centre **Time period:** between the years 2012 and 2015 **Country study is conducted:** UK
Index tests	**Name index test:** MRD **Details of conducting index test:** MRD examinations were performed on a 1.5 T closed magnet (MAGNETOM Symphony, Siemens, Germany). Participants lay supine on the MRI table on a waterproof mat, knees slightly flexed with a pillow underneath, and legs slightly apart. A flexible transmit/receive radiofrequency Siemens 6 channel multiphase coil was wrapped around the pelvis **Imaging acquisition:** The MRD protocol comprised T2-weighted (T2W) spin echo sagittal and T1-weighted spin echo axial sequences through the pelvis. T2W spin echo high-resolution oblique axial images perpendicular to the vagina and through the puborectalis sling and oblique coronal sequences parallel to the vagina through the puborectalis sling were taken to assess pelvic floor morphology at rest. Balanced steady-state free procession sequence (TrueFISP) was used to assess dynamic pelvic floor function. A dynamic True-FISP coronal squeeze for 5 seconds followed immediately by a bear down was repeated twice in the same scan acquisition – angle parallel with vagina and placed mid-rectum. A dynamic TrueFISP in the mid-sagittal plane during bear down and coronal view was obtained through the bladder base again during bear down. GE Polaris II Ultrasound Gel (120

Zafar 2017 *(Continued)*

	ml) was then inserted into the rectum. The participant was asked to hold onto the gel and lie on their back for a couple of minutes before being returned into the scanner and a new localiser obtained. The participant was told to perform a continuous push down of about 12 – 15 seconds. If the participant was not successful in evacuating the gel at first attempt, then 2 further attempts were allowed **Imaging analysis:** The EP and MRD scans were reported by 2 consultant radiologists with a special interest in gastrointestinal imaging and considerable experience in pelvic floor imaging **Threshold test positivity:** Rectocele > 20 mm; intussusception circumferential full thickness
Target condition and reference standard(s)	**Name index test 'EP':** Evacuation proctography (EP) **Details of conducting evacuation proctography:** Participants were seated on a commode, placed on the footrest of an upright-positioned examination table in front of a fluoroscopic unit. Thickened barium paste was instilled with the participant in a lateral decubitus position using a Foley's catheter. The tube was removed and the participant was asked to sit on the modified commode **Imaging acquisition:** The images were obtained with the participant at rest and attempting to defaecate **Imaging analysis:** The EP and MRD scans were reported by 2 consultant radiologists with a special interest in gastrointestinal imaging and considerable experience in pelvic floor imaging **Threshold test positivity:** Rectocele > 20 mm; intussusception circumferential full thickness
Flow and timing	**Enrolment and exclusions (+ reasons):** All were included in the 2 x 2 tables. **Nr analysed:** 55 **Time interval (+ interventions) between index test and reference standard:** The tests were performed at least 2 weeks apart and in no particular order
Comparative	Not applicable
Notes	

Methodological quality

Item	Authors' judgement	Risk of bias	Applicability concerns
DOMAIN 1: Patient Selection			
Was a consecutive or random sample of patients enrolled?	Yes		
Did the study avoid inappropriate exclusions?	Yes		
Could the selection of patients have introduced bias?		Low risk	
Are the included patients only female or are test accuracy data provided for only female participants?			
Do the included patients only have ODS symptoms?			
Are there concerns that the included patients and			Low concern

Zafar 2017 *(Continued)*

setting do not match the review question?			
DOMAIN 2: Index Test (MRI or Ultrasound)			
Was the threshold for test positivity pre-specified?	Yes		
Where the index test results interpreted without knowledge of the results of the other index test(s)?	Yes		
Could the conduct or interpretation of the index test have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
For MRI was a scanner used with Tesla 1 or higher?			
Are there concerns that the index test, its conduct, or interpretation differ from the review question?			Low concern
DOMAIN 3: Reference Standard			
Was the threshold for test positivity pre-specified?	Yes		
Where the EP results interpreted without the knowledge of the results of the other index test(s)?	Yes		
Could the reference standard, its conduct, or its interpretation have introduced bias?		Low risk	
If a reference line was used, was it the PCL?			
Are there concerns that the target condition as defined by the reference standard does not match the question?			Low concern
DOMAIN 4: Flow and Timing			
Was there an appropriate interval between index test and reference standard?	Yes		
Did all patients receive the same reference standard?	Yes		

Zafar 2017 *(Continued)*

Were all patients included in the analysis?	Yes
Could the patient flow have introduced bias?	Low risk

In the domain 'reference standard' the results for EP are presented. Note that in this review EP is considered as index test and not as reference standard.
ARA: anorectal angle; ARJ: ano-rectal junction; CCD: colpo-cysto-defaecation; DAE: dynamic endosonography; DTPU: dynamic transperineal ultrasound; DEP: dynamic evacuation proctography; EP: evacuation proctography; EVUS: endovaginal ultrasound; FOV: field of view; MRD: magnetic resonance defaecography; N/A: not applicable; ODS: obstructed defaecation syndrome; PCL: pubococcygeal line; TPUS: transperineal ultrasound;

Characteristics of excluded studies *[ordered by study ID]*

Study	Reason for exclusion
Beer-Gabel 2002	Not correct outcome measure: Accuracy of measurements of target conditions Tests accuracy data requested from authors, but data not available
Beer-Gabel 2010	Not correct outcome measure: Association between IBS and ODS Tests accuracy data requested from authors, but data not available
Beer-Gabel 2011	Not correct outcome measure: incidence and type of IBS-related symptomatology in women with ODS Tests accuracy data requested from authors, but data not available
Bot-Robin 2011	Not correct outcome measure: Feasibility of a surgical concomitant treatment of a rectal and pelvic prolapse with a mesh sutured to the rectum during a vaginal approach Tests accuracy data requested from authors, but data not available
Bussen 2003	Not correct outcome measure: accuracy of measurements of target conditions Tests accuracy data requested from authors, but data not available.
Cappabianca 2011	Case-control study design: Only participants with enterocele on EP underwent MRI (index test)
Cerdán 2011	Not correct outcome measure: Analysis of functional and post-operative results in participants who underwent surgery for enterocele Test accuracy data requested from authors, but no reply received
Chatoor 2007	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Chung 2003	More men than women included. Patients < 18 years included. Requested data on women > 18 years only, but data not available
Dekel 2015	Not correct outcome measure: To evaluate the value of balloon expulsion test in the diagnostic process of pelvic dyssynergia Test accuracy data requested from authors, but no reply received

Study	Reason for exclusion
Deval 2003	Did not use evacuation proctography. Verified with authors
Dvorkin 2004	Case-control study design: Only participants with intussusception on EP underwent MRI (index test)
Ferrari 2019	Not correct outcome measure: to assess characteristics of participants with primary symptoms of faecal incontinence in a tertiary referral centre Test accuracy data requested from authors, but no reply received.
Fletcher 2003	No evacuation proctography: Scintinography was used as reference standard (nuclear instead of x-ray)
Goffredo 2010	Case-control study design: only participants with anismus on US underwent EP (reference standard)
Groenendijk 2009	Not correct outcome measure: to establish the effects of additional diagnostic tests compared to a consensus outcome on treatment selection in primary pelvic organ prolapse Tests accuracy data requested from authors, but data not available
Healy 1998	Case-control study design: Only participants with normal EP underwent MRI (index test)
Imanova 2017	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Kaufman 2001	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Kawata 2010	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Köhler 2012	Not correct outcome measure: to investigate the long-term results and predictive factors for outcome after STARR procedure Test accuracy data requested from authors, but no reply received
Mege 2013	Not correct outcome measure: To identify predictive factors for long-term symptomatic failure following elytrocele surgical correction by abdominal approach Test accuracy data requested from authors and received. Selected participant population: Only women with symptomatic elytrocele
Ortega 2011	Not correct outcome measure: To analyse functional results on participants who underwent surgery for enterocele Test accuracy data requested from authors, but no reply received
Otto 2011	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Pannu 2009	Not able to extract all test accuracy data Test accuracy data requested from authors, but no reply received

Study	Reason for exclusion
Pescatori 2006	Not correct outcome measure: To evaluate occult disorders in participants undergoing surgical treatment for ODS Tests accuracy data requested from authors, but data not available
Pescatori 2009a	Not correct outcome measure: to assess participants following performance of the STARR procedure for ODS where the procedure was complicated or had failed Tests accuracy data requested from authors, but data not available
Pescatori 2009b	Not correct outcome measure: to investigate the results of an abdominoperineal procedure aimed at treating enterorectocele with recto-rectal intussusception in 1 stage Tests accuracy data requested from authors, but data not available
Petersen 2006	Not correct outcome measure: to evaluate a combined procedure of transanal rectal resection with a simultaneous laparoscopy for participants with obstructed defaecation syndrome and an enterocele Test accuracy data requested from authors, but no reply received
Renzi 2016	Not correct outcome measure: to report the short-term preliminary results of a novel surgical procedure, transverse perineal support, for the correction of pathological perineal descent Test accuracy data requested from authors, but no reply received
Ricchiuti 2016	Not correct outcome measure: To evaluate if body position affects the assessment of puborectalis muscle length (PRL) and anorectal angle (ARA. Test accuracy data requested from authors, but no reply received
Rizal 2014	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Ron 2018	Not correct outcome measure: to assess the value of specially-designed toilet seat for participants suffering from obstructed defaecation type of constipation Test accuracy data requested from authors, but no reply received
Schoenenberger 1998	Gender of population unknown; both men and women included. Data on women only requested from the authors, but data not available. Female/male ratio approximately 2:1
Song 2009	Not able to extract test accuracy data Test accuracy data requested from authors, but no reply received
Tsar'kov 2012	Not correct outcome measure: To evaluate in complex the effectiveness of transvaginal mesh implants in women with obstructed defaecation (OD) syndrome based on the comparison of preoperative and postoperative results Test accuracy data requested from authors, but no reply received
Wang 2005	Not able to extract all test accuracy data Test accuracy data requested from authors, but no reply received
Xiong 2006	Case-control study design: Assessing imaging in women with and without anismus
Zeng 2019	Not able to extract test accuracy data

Study	Reason for exclusion
	Test accuracy data requested from authors, but no reply received

DATA

Presented below are all the data for all of the tests entered into the review.

Table Tests. Data tables by test

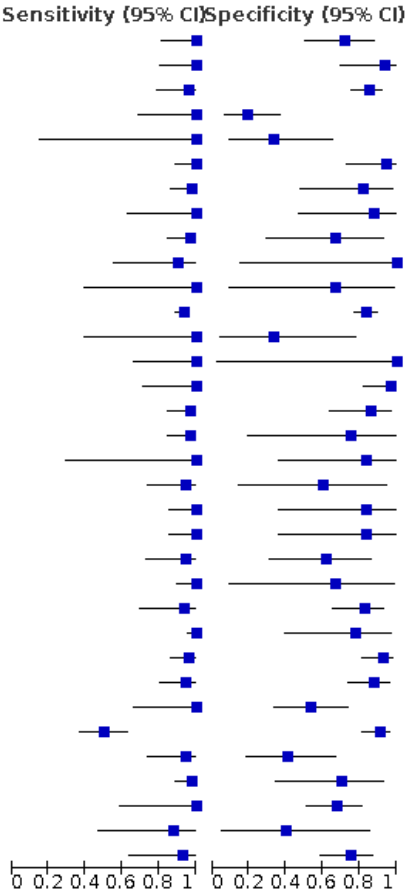
Test	No. of studies	No. of participants
1 EP - Rectocele - LCA	34	1737
2 EP - Enterocele - LCA	31	2233
3 EP - Intussusception - LCA	27	1613
4 EP - Anismus - LCA	15	985
5 EP - PFD - LCA	10	476
6 MRI - Rectocele - LCA	19	659
7 MRI - Enterocele - LCA	17	1222
8 MRI - Intussusception - LCA	12	536
9 MRI - Anismus - LCA	7	287
10 MRI - PFD - LCA	7	350
11 TPUS - Rectocele - LCA	11	988
12 TPUS - Enterocele - LCA	10	963
13 TPUS - Intussusception - LCA	10	664
14 TPUS - Anismus - LCA	5	651
15 TPUS - PFD - LCA	1	54
16 EVUS - Rectocele - LCA	2	454
17 EVUS - Enterocele - LCA	3	471
18 EVUS - Intussusception - LCA	2	454
19 EVUS - Anismus - LCA	2	454
20 DAE - Rectocele - LCA	2	99
21 DAE - Enterocele - LCA	2	70

Test	No. of studies	No. of participants
22 DAE - Intussusception - LCA	2	99
23 DAE - PFD - LCA	2	99
24 EDF - Rectocele - LCA	4	169
25 EDF - Enterocele - LCA	3	139
26 EDF - Intussusception - LCA	4	169
27 EDF - Anismus - LCA	4	169
28 EDF - PFD - LCA	1	29

Test 1. EP - Rectocele - LCA

EP - Rectocele - LCA

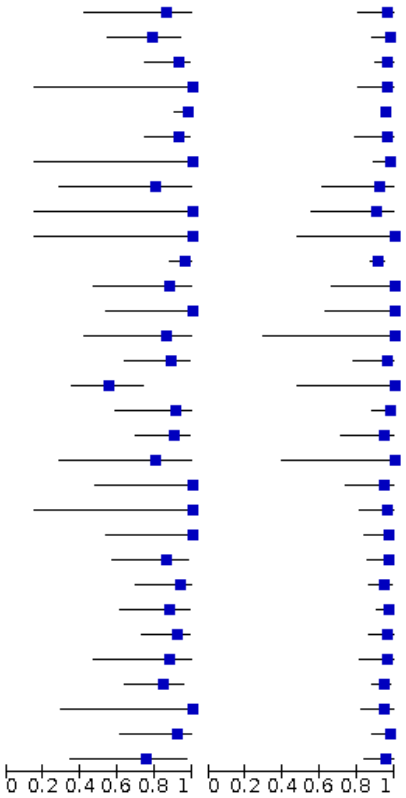
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	18	7	0	18	1.00 [0.81, 1.00]	0.72 [0.51, 0.88]
Beer-Gabel 2004	17	1	0	15	1.00 [0.80, 1.00]	0.94 [0.70, 1.00]
Beer-Gabel 2015	23	12	1	69	0.96 [0.79, 1.00]	0.85 [0.76, 0.92]
Brusciano 2007	10	25	0	6	1.00 [0.69, 1.00]	0.19 [0.07, 0.37]
Dellemare 1994	2	8	0	4	1.00 [0.16, 1.00]	0.33 [0.10, 0.65]
Faucheron 2014	32	1	0	17	1.00 [0.89, 1.00]	0.94 [0.73, 1.00]
Fiaschetti 2013	37	2	1	9	0.97 [0.86, 1.00]	0.82 [0.48, 0.98]
Foti 2013	8	1	0	7	1.00 [0.63, 1.00]	0.88 [0.47, 1.00]
Grasso 2007	33	3	1	6	0.97 [0.85, 1.00]	0.67 [0.30, 0.93]
Gufler 1999	9	0	1	2	0.90 [0.55, 1.00]	1.00 [0.16, 1.00]
Gufler 2004	4	1	0	2	1.00 [0.40, 1.00]	0.67 [0.09, 0.99]
Hainsworth 2016	164	24	12	124	0.93 [0.88, 0.96]	0.84 [0.77, 0.89]
Healy 1997	4	4	0	2	1.00 [0.40, 1.00]	0.33 [0.04, 0.78]
Kelvin 2000	9	0	0	1	1.00 [0.66, 1.00]	1.00 [0.03, 1.00]
Lienemann 1997	11	1	0	28	1.00 [0.72, 1.00]	0.97 [0.82, 1.00]
Martellucci 2011	32	3	1	18	0.97 [0.84, 1.00]	0.86 [0.64, 0.97]
Martin 2017	33	1	1	3	0.97 [0.85, 1.00]	0.75 [0.19, 0.99]
Matsuoka 2000	3	1	0	5	1.00 [0.29, 1.00]	0.83 [0.36, 1.00]
Miravalle 2016	18	2	1	3	0.95 [0.74, 1.00]	0.60 [0.15, 0.95]
Murad-Regadas 2008	24	1	0	5	1.00 [0.86, 1.00]	0.83 [0.36, 1.00]
Murad-Regadas 2011	23	1	0	5	1.00 [0.85, 1.00]	0.83 [0.36, 1.00]
Perniola 2008	17	5	1	8	0.94 [0.73, 1.00]	0.62 [0.32, 0.86]
Pilkington 2012	35	1	0	2	1.00 [0.90, 1.00]	0.67 [0.09, 0.99]
Poncelet 2017	15	6	1	28	0.94 [0.70, 1.00]	0.82 [0.65, 0.93]
Regadas 2011	77	2	0	7	1.00 [0.95, 1.00]	0.78 [0.40, 0.97]
Ron 2012	48	4	2	48	0.96 [0.86, 1.00]	0.92 [0.81, 0.98]
Steensma 2010	32	5	2	36	0.94 [0.80, 0.99]	0.88 [0.74, 0.96]
Vanbeckevoort 1999	9	12	0	14	1.00 [0.66, 1.00]	0.54 [0.33, 0.73]
Van Gruting 2017	32	6	32	61	0.50 [0.37, 0.63]	0.91 [0.82, 0.97]
Van Iersel 2017	18	10	1	7	0.95 [0.74, 1.00]	0.41 [0.18, 0.67]
Vitton 2011	45	3	1	7	0.98 [0.88, 1.00]	0.70 [0.35, 0.93]
Weemhoff 2013	7	14	0	29	1.00 [0.59, 1.00]	0.67 [0.51, 0.81]
Zafar 2012	7	3	1	2	0.88 [0.47, 1.00]	0.40 [0.05, 0.85]
Zafar 2017	12	10	1	30	0.92 [0.64, 1.00]	0.75 [0.59, 0.87]



Test 2. EP - Enterocele - LCA

EP - Enterocele - LCA

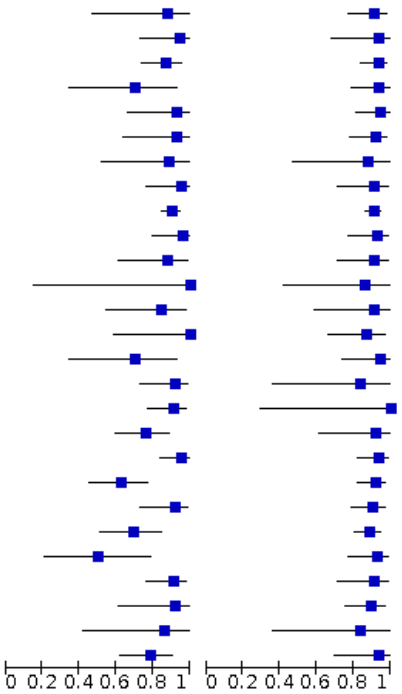
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	6	1	1	25	0.86 [0.42, 1.00]	0.96 [0.80, 1.00]
Beer-Gabel 2008	15	1	4	42	0.79 [0.54, 0.94]	0.98 [0.88, 1.00]
Beer-Gabel 2015	24	3	2	76	0.92 [0.75, 0.99]	0.96 [0.89, 0.99]
Brusciano 2007	2	1	0	25	1.00 [0.16, 1.00]	0.96 [0.80, 1.00]
Faggian 2013	74	26	2	512	0.97 [0.91, 1.00]	0.95 [0.93, 0.97]
Faucheron 2014	24	1	2	23	0.92 [0.75, 0.99]	0.96 [0.79, 1.00]
Fiaschetti 2013	2	1	0	46	1.00 [0.16, 1.00]	0.98 [0.89, 1.00]
Foti 2013	4	1	1	11	0.80 [0.28, 0.99]	0.92 [0.62, 1.00]
Gufler 1999	2	1	0	9	1.00 [0.16, 1.00]	0.90 [0.55, 1.00]
Gufler 2004	2	0	0	5	1.00 [0.16, 1.00]	1.00 [0.48, 1.00]
Hainsworth 2016	54	23	2	244	0.96 [0.88, 1.00]	0.91 [0.87, 0.94]
Halligan 1996	7	0	1	9	0.88 [0.47, 1.00]	1.00 [0.66, 1.00]
Karaus 2000	6	0	0	8	1.00 [0.54, 1.00]	1.00 [0.63, 1.00]
Kelvin 2000	6	0	1	3	0.86 [0.42, 1.00]	1.00 [0.29, 1.00]
Lienemann 1997	15	1	2	22	0.88 [0.64, 0.99]	0.96 [0.78, 1.00]
Lienemann 2000	16	0	13	5	0.55 [0.36, 0.74]	1.00 [0.48, 1.00]
Martellucci 2011	10	1	1	42	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]
Martin 2017	19	1	2	16	0.90 [0.70, 0.99]	0.94 [0.71, 1.00]
Matsuoka 2000	4	0	1	4	0.80 [0.28, 0.99]	1.00 [0.40, 1.00]
Miravalle 2016	5	1	0	18	1.00 [0.48, 1.00]	0.95 [0.74, 1.00]
Murad-Regadas 2011	2	1	0	26	1.00 [0.16, 1.00]	0.96 [0.81, 1.00]
Pilkington 2012	6	1	0	31	1.00 [0.54, 1.00]	0.97 [0.84, 1.00]
Poncelet 2017	12	1	2	35	0.86 [0.57, 0.98]	0.97 [0.85, 1.00]
Regadas 2011	15	4	1	66	0.94 [0.70, 1.00]	0.94 [0.86, 0.98]
Ron 2012	14	3	2	83	0.88 [0.62, 0.98]	0.97 [0.90, 0.99]
Steensma 2010	22	2	2	49	0.92 [0.73, 0.99]	0.96 [0.87, 1.00]
Vanbeckevoort 1999	7	1	1	26	0.88 [0.47, 1.00]	0.96 [0.81, 1.00]
Van Gruting 2017	21	6	4	100	0.84 [0.64, 0.95]	0.94 [0.88, 0.98]
Van Iersel 2017	3	2	0	36	1.00 [0.29, 1.00]	0.95 [0.82, 0.99]
Vitton 2011	11	1	1	43	0.92 [0.62, 1.00]	0.98 [0.88, 1.00]
Weemhoff 2013	6	2	2	40	0.75 [0.35, 0.97]	0.95 [0.84, 0.99]



Test 3. EP - Intussusception - LCA

EP - Intussusception - LCA

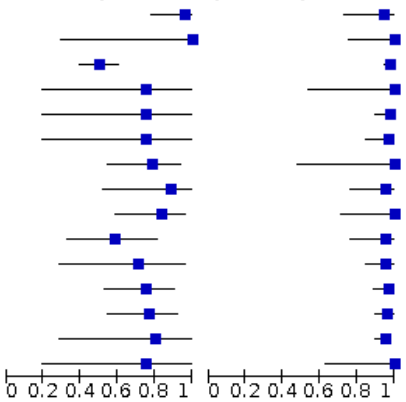
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	7	3	1	32	0.88 [0.47, 1.00]	0.91 [0.77, 0.98]
Beer-Gabel 2004	17	1	1	14	0.94 [0.73, 1.00]	0.93 [0.68, 1.00]
Beer-Gabel 2015	40	4	6	55	0.87 [0.74, 0.95]	0.93 [0.84, 0.98]
Brusciano 2007	7	2	3	29	0.70 [0.35, 0.93]	0.94 [0.79, 0.99]
Faucheron 2014	13	2	1	34	0.93 [0.66, 1.00]	0.94 [0.81, 0.99]
Fiaschetti 2013	12	3	1	33	0.92 [0.64, 1.00]	0.92 [0.78, 0.98]
Foti 2013	8	1	1	7	0.89 [0.52, 1.00]	0.88 [0.47, 1.00]
Grasso 2007	20	2	1	20	0.95 [0.76, 1.00]	0.91 [0.71, 0.99]
Hainsworth 2016	125	16	13	169	0.91 [0.84, 0.95]	0.91 [0.86, 0.95]
Martellucci 2011	24	2	1	27	0.96 [0.80, 1.00]	0.93 [0.77, 0.99]
Martin 2017	14	2	2	20	0.88 [0.62, 0.98]	0.91 [0.71, 0.99]
Matsuoka 2000	2	1	0	6	1.00 [0.16, 1.00]	0.86 [0.42, 1.00]
Miravalle 2016	11	1	2	10	0.85 [0.55, 0.98]	0.91 [0.59, 1.00]
Murad-Regadas 2008	7	3	0	20	1.00 [0.59, 1.00]	0.87 [0.66, 0.97]
Murad-Regadas 2011	7	1	3	18	0.70 [0.35, 0.93]	0.95 [0.74, 1.00]
Perniola 2008	22	1	2	5	0.92 [0.73, 0.99]	0.83 [0.36, 1.00]
Pilkington 2012	32	0	3	3	0.91 [0.77, 0.98]	1.00 [0.29, 1.00]
Poncelet 2017	29	1	9	11	0.76 [0.60, 0.89]	0.92 [0.62, 1.00]
Regadas 2011	39	3	2	42	0.95 [0.83, 0.99]	0.93 [0.82, 0.99]
Ron 2012	25	5	15	57	0.63 [0.46, 0.77]	0.92 [0.82, 0.97]
Steensma 2010	22	5	2	46	0.92 [0.73, 0.99]	0.90 [0.79, 0.97]
Van Gruting 2017	23	11	10	87	0.70 [0.51, 0.84]	0.89 [0.81, 0.94]
Van Iersel 2017	6	2	6	27	0.50 [0.21, 0.79]	0.93 [0.77, 0.99]
Vitton 2011	31	2	3	20	0.91 [0.76, 0.98]	0.91 [0.71, 0.99]
Weemhoff 2013	11	4	1	34	0.92 [0.62, 1.00]	0.89 [0.75, 0.97]
Zafar 2012	6	1	1	5	0.86 [0.42, 1.00]	0.83 [0.36, 1.00]
Zafar 2017	29	1	8	15	0.78 [0.62, 0.90]	0.94 [0.70, 1.00]



Test 4. EP - Anismus - LCA

EP - Anismus - LCA

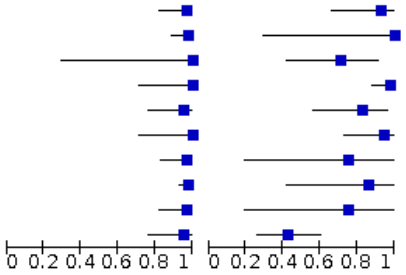
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Brusciano 2007	22	1	1	17	0.96 [0.78, 1.00]	0.94 [0.73, 1.00]
Foti 2013	3	0	0	13	1.00 [0.29, 1.00]	1.00 [0.75, 1.00]
Hainsworth 2016	46	6	46	226	0.50 [0.39, 0.61]	0.97 [0.94, 0.99]
Healy 1997	3	0	1	6	0.75 [0.19, 0.99]	1.00 [0.54, 1.00]
Martellucci 2011	3	1	1	49	0.75 [0.19, 0.99]	0.98 [0.89, 1.00]
Martin 2017	3	1	1	33	0.75 [0.19, 0.99]	0.97 [0.85, 1.00]
Miravalle 2016	15	0	4	5	0.79 [0.54, 0.94]	1.00 [0.48, 1.00]
Murad-Regadas 2008	8	1	1	20	0.89 [0.52, 1.00]	0.95 [0.76, 1.00]
Murad-Regadas 2011	15	0	3	11	0.83 [0.59, 0.96]	1.00 [0.72, 1.00]
Pilkington 2012	10	1	7	20	0.59 [0.33, 0.82]	0.95 [0.76, 1.00]
Poncelet 2017	5	2	2	41	0.71 [0.29, 0.96]	0.95 [0.84, 0.99]
Regadas 2011	18	2	6	60	0.75 [0.53, 0.90]	0.97 [0.89, 1.00]
Ron 2012	17	3	5	77	0.77 [0.55, 0.92]	0.96 [0.89, 0.99]
Van Gruting 2017	4	6	1	120	0.80 [0.28, 0.99]	0.95 [0.90, 0.98]
Zafar 2012	3	0	1	8	0.75 [0.19, 0.99]	1.00 [0.63, 1.00]



Test 5. EP - PFD - LCA

EP - PFD - LCA

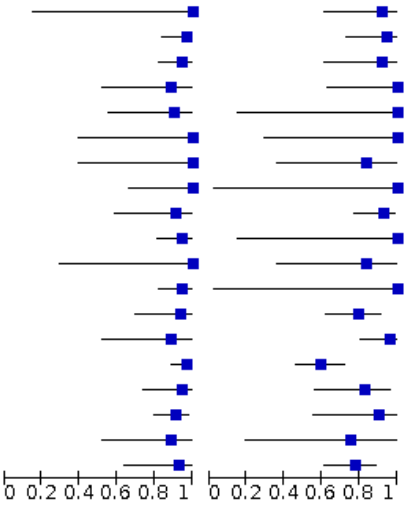
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	28	1	1	13	0.97 [0.82, 1.00]	0.93 [0.66, 1.00]
Fiaschetti 2013	45	0	1	3	0.98 [0.88, 1.00]	1.00 [0.29, 1.00]
Foti 2013	3	4	0	10	1.00 [0.29, 1.00]	0.71 [0.42, 0.92]
Martellucci 2011	11	1	0	42	1.00 [0.72, 1.00]	0.98 [0.88, 1.00]
Martin 2017	20	3	1	14	0.95 [0.76, 1.00]	0.82 [0.57, 0.96]
Murad-Regadas 2011	11	1	0	17	1.00 [0.72, 1.00]	0.94 [0.73, 1.00]
Vanbeckevoort 1999	30	1	1	3	0.97 [0.83, 1.00]	0.75 [0.19, 0.99]
Van Gruting 2017	112	1	3	6	0.97 [0.93, 0.99]	0.86 [0.42, 1.00]
Van Iersel 2017	28	1	1	3	0.97 [0.82, 1.00]	0.75 [0.19, 0.99]
Vitton 2011	20	20	1	15	0.95 [0.76, 1.00]	0.43 [0.26, 0.61]



Test 6. MRI - Rectocele - LCA

MRI - Rectocele - LCA

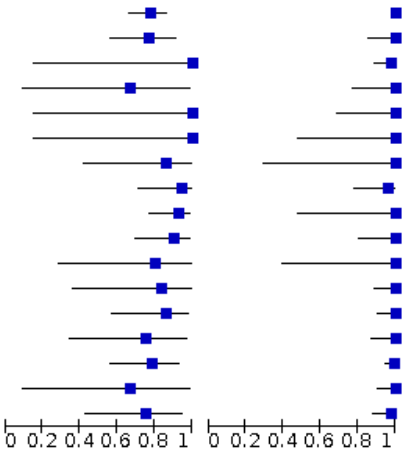
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Dellemare 1994	2	1	0	11	1.00 [0.16, 1.00]	0.92 [0.62, 1.00]
Faucheron 2014	31	1	1	17	0.97 [0.84, 1.00]	0.94 [0.73, 1.00]
Fiaschetti 2013	35	1	2	11	0.95 [0.82, 0.99]	0.92 [0.62, 1.00]
Foti 2013	8	0	1	8	0.89 [0.52, 1.00]	1.00 [0.63, 1.00]
Gufler 1999	9	0	1	2	0.90 [0.55, 1.00]	1.00 [0.16, 1.00]
Gufler 2004	4	0	0	3	1.00 [0.40, 1.00]	1.00 [0.29, 1.00]
Healy 1997	4	1	0	5	1.00 [0.40, 1.00]	0.83 [0.36, 1.00]
Kelvin 2000	9	0	0	1	1.00 [0.66, 1.00]	1.00 [0.03, 1.00]
Lienemann 1997	10	2	1	27	0.91 [0.59, 1.00]	0.93 [0.77, 0.99]
Martin 2017	34	0	2	2	0.94 [0.81, 0.99]	1.00 [0.16, 1.00]
Matsuoka 2000	3	1	0	5	1.00 [0.29, 1.00]	0.83 [0.36, 1.00]
Pilkington 2012	35	0	2	1	0.95 [0.82, 0.99]	1.00 [0.03, 1.00]
Poncelet 2017	15	7	1	27	0.94 [0.70, 1.00]	0.79 [0.62, 0.91]
Vanbeckevoort 1999	8	1	1	25	0.89 [0.52, 1.00]	0.96 [0.80, 1.00]
Van Gruting 2017	58	25	2	37	0.97 [0.88, 1.00]	0.60 [0.46, 0.72]
Van Iersel 2017	18	3	1	14	0.95 [0.74, 1.00]	0.82 [0.57, 0.96]
Vitton 2011	42	1	4	9	0.91 [0.79, 0.98]	0.90 [0.55, 1.00]
Zafar 2012	8	1	1	3	0.89 [0.52, 1.00]	0.75 [0.19, 0.99]
Zafar 2017	12	9	1	31	0.92 [0.64, 1.00]	0.78 [0.62, 0.89]



Test 7. MRI - Enterocele - LCA

MRI - Enterocele - LCA

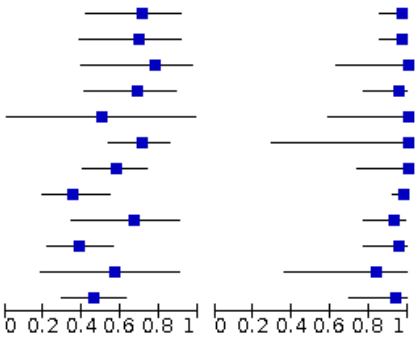
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Faggian 2013	58	1	17	538	0.77 [0.66, 0.86]	1.00 [0.99, 1.00]
Faucheron 2014	20	0	6	24	0.77 [0.56, 0.91]	1.00 [0.86, 1.00]
Fiaschetti 2013	2	1	0	46	1.00 [0.16, 1.00]	0.98 [0.89, 1.00]
Foti 2013	2	0	1	14	0.67 [0.09, 0.99]	1.00 [0.77, 1.00]
Gufler 1999	2	0	0	10	1.00 [0.16, 1.00]	1.00 [0.69, 1.00]
Gufler 2004	2	0	0	5	1.00 [0.16, 1.00]	1.00 [0.48, 1.00]
Kelvin 2000	6	0	1	3	0.86 [0.42, 1.00]	1.00 [0.29, 1.00]
Lienemann 1997	16	1	1	22	0.94 [0.71, 1.00]	0.96 [0.78, 1.00]
Lienemann 2000	27	0	2	5	0.93 [0.77, 0.99]	1.00 [0.48, 1.00]
Martin 2017	19	0	2	17	0.90 [0.70, 0.99]	1.00 [0.80, 1.00]
Matsuoka 2000	4	0	1	4	0.80 [0.28, 0.99]	1.00 [0.40, 1.00]
Pilkington 2012	5	0	1	32	0.83 [0.36, 1.00]	1.00 [0.89, 1.00]
Poncelet 2017	12	0	2	36	0.86 [0.57, 0.98]	1.00 [0.90, 1.00]
Vanbeckevoort 1999	6	0	2	27	0.75 [0.35, 0.97]	1.00 [0.87, 1.00]
Van Gruting 2017	18	1	5	98	0.78 [0.56, 0.93]	0.99 [0.95, 1.00]
Van Iersel 2017	2	0	1	38	0.67 [0.09, 0.99]	1.00 [0.91, 1.00]
Vitton 2011	9	1	3	43	0.75 [0.43, 0.95]	0.98 [0.88, 1.00]



Test 8. MRI - Intussusception - LCA

MRI - Intussusception - LCA

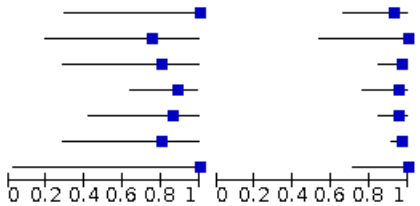
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Faucheron 2014	10	1	4	35	0.71 [0.42, 0.92]	0.97 [0.85, 1.00]
Fiaschetti 2013	9	1	4	35	0.69 [0.39, 0.91]	0.97 [0.85, 1.00]
Foti 2013	7	0	2	8	0.78 [0.40, 0.97]	1.00 [0.63, 1.00]
Martin 2017	11	1	5	21	0.69 [0.41, 0.89]	0.95 [0.77, 1.00]
Matsuoka 2000	1	0	1	7	0.50 [0.01, 0.99]	1.00 [0.59, 1.00]
Pilkington 2012	25	0	10	3	0.71 [0.54, 0.85]	1.00 [0.29, 1.00]
Poncelet 2017	22	0	16	12	0.58 [0.41, 0.74]	1.00 [0.74, 1.00]
Van Gruting 2017	11	2	20	89	0.35 [0.19, 0.55]	0.98 [0.92, 1.00]
Van Iersel 2017	8	2	4	27	0.67 [0.35, 0.90]	0.93 [0.77, 0.99]
Vitton 2011	13	1	21	21	0.38 [0.22, 0.56]	0.95 [0.77, 1.00]
Zafar 2012	4	1	3	5	0.57 [0.18, 0.90]	0.83 [0.36, 1.00]
Zafar 2017	17	1	20	15	0.46 [0.29, 0.63]	0.94 [0.70, 1.00]



Test 9. MRI - Anismus - LCA

MRI - Anismus - LCA

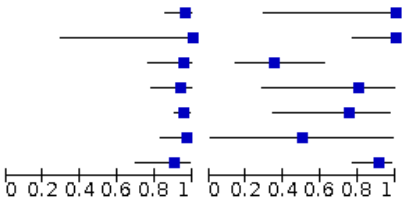
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Foti 2013	3	1	0	13	1.00 [0.29, 1.00]	0.93 [0.66, 1.00]
Healy 1997	3	0	1	6	0.75 [0.19, 0.99]	1.00 [0.54, 1.00]
Martin 2017	4	1	1	32	0.80 [0.28, 0.99]	0.97 [0.84, 1.00]
Pilkington 2012	15	1	2	20	0.88 [0.64, 0.99]	0.95 [0.76, 1.00]
Poncelet 2017	6	2	1	41	0.86 [0.42, 1.00]	0.95 [0.84, 0.99]
Van Gruting 2017	4	4	1	113	0.80 [0.28, 0.99]	0.97 [0.91, 0.99]
Zafar 2012	1	0	0	11	1.00 [0.03, 1.00]	1.00 [0.72, 1.00]



Test 10. MRI - PFD - LCA

MRI - PFD - LCA

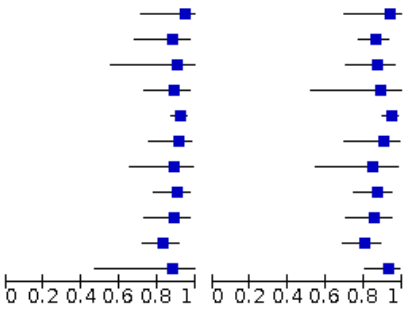
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Fiaschetti 2013	44	0	2	3	0.96 [0.85, 0.99]	1.00 [0.29, 1.00]
Foti 2013	3	0	0	14	1.00 [0.29, 1.00]	1.00 [0.77, 1.00]
Martin 2017	20	11	1	6	0.95 [0.76, 1.00]	0.35 [0.14, 0.62]
Vanbeckevoort 1999	28	1	2	4	0.93 [0.78, 0.99]	0.80 [0.28, 0.99]
Van Gruting 2017	109	2	5	6	0.96 [0.90, 0.99]	0.75 [0.35, 0.97]
Van Iersel 2017	30	1	1	1	0.97 [0.83, 1.00]	0.50 [0.01, 0.99]
Vitton 2011	19	3	2	32	0.90 [0.70, 0.99]	0.91 [0.77, 0.98]



Test 11. TPUS - Rectocele - LCA

TPUS - Rectocele - LCA

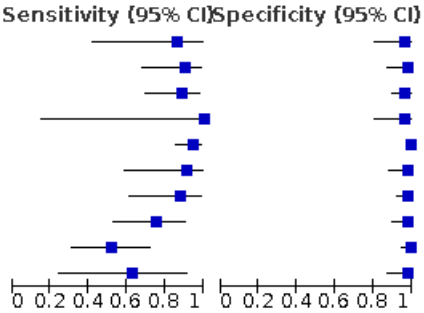
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	16	1	1	15	0.94 [0.71, 1.00]	0.94 [0.70, 1.00]
Beer-Gabel 2015	21	11	3	70	0.88 [0.68, 0.97]	0.86 [0.77, 0.93]
Brusciano 2007	9	4	1	27	0.90 [0.55, 1.00]	0.87 [0.70, 0.96]
Grasso 2007	30	1	4	8	0.88 [0.73, 0.97]	0.89 [0.52, 1.00]
Hainsworth 2016	161	8	14	140	0.92 [0.87, 0.96]	0.95 [0.90, 0.98]
Martellucci 2011	30	2	3	19	0.91 [0.76, 0.98]	0.90 [0.70, 0.99]
Perniola 2008	16	2	2	11	0.89 [0.65, 0.99]	0.85 [0.55, 0.98]
Ron 2012	45	7	5	45	0.90 [0.78, 0.97]	0.87 [0.74, 0.94]
Steensma 2010	30	6	4	35	0.88 [0.73, 0.97]	0.85 [0.71, 0.94]
Van Gruting 2017	54	13	11	53	0.83 [0.72, 0.91]	0.80 [0.69, 0.89]
Weemhoff 2013	7	3	1	39	0.88 [0.47, 1.00]	0.93 [0.81, 0.99]



Test 12. TPUS - Enterocele - LCA

TPUS - Enterocele - LCA

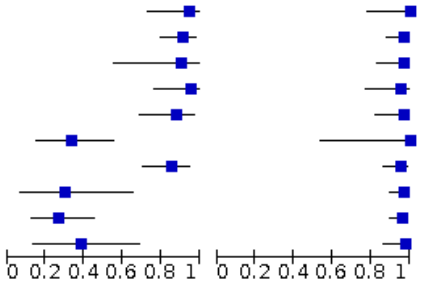
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	6	1	1	25	0.86 [0.42, 1.00]	0.96 [0.80, 1.00]
Beer-Gabel 2008	18	1	2	41	0.90 [0.68, 0.99]	0.98 [0.87, 1.00]
Beer-Gabel 2015	23	3	3	76	0.88 [0.70, 0.98]	0.96 [0.89, 0.99]
Brusciano 2007	2	1	0	25	1.00 [0.16, 1.00]	0.96 [0.80, 1.00]
Hainsworth 2016	53	2	3	265	0.95 [0.85, 0.99]	0.99 [0.97, 1.00]
Martellucci 2011	10	1	1	42	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]
Ron 2012	14	2	2	84	0.88 [0.62, 0.98]	0.98 [0.92, 1.00]
Steensma 2010	18	1	6	50	0.75 [0.53, 0.90]	0.98 [0.90, 1.00]
Van Gruting 2017	13	1	12	105	0.52 [0.31, 0.72]	0.99 [0.95, 1.00]
Weemhoff 2013	5	1	3	41	0.63 [0.24, 0.91]	0.98 [0.87, 1.00]



Test 13. TPUS - Intussusception - LCA

TPUS - Intussusception - LCA

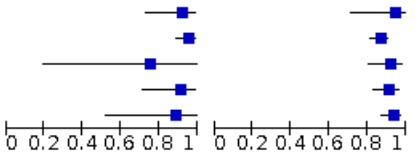
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Beer-Gabel 2004	17	0	1	15	0.94 [0.73, 1.00]	1.00 [0.78, 1.00]		
Beer-Gabel 2015	42	2	4	57	0.91 [0.79, 0.98]	0.97 [0.88, 1.00]		
Brusciano 2007	9	1	1	30	0.90 [0.55, 1.00]	0.97 [0.83, 1.00]		
Grasso 2007	20	1	1	21	0.95 [0.76, 1.00]	0.95 [0.77, 1.00]		
Martellucci 2011	22	1	3	28	0.88 [0.69, 0.97]	0.97 [0.82, 1.00]		
Perniola 2008	8	0	16	6	0.33 [0.16, 0.55]	1.00 [0.54, 1.00]		
Ron 2012	34	3	6	59	0.85 [0.70, 0.94]	0.95 [0.87, 0.99]		
Steensma 2010	3	2	7	63	0.30 [0.07, 0.65]	0.97 [0.89, 1.00]		
Van Gruting 2017	9	4	24	94	0.27 [0.13, 0.46]	0.96 [0.90, 0.99]		
Weemhoff 2013	5	1	8	36	0.38 [0.14, 0.68]	0.97 [0.86, 1.00]		



Test 14. TPUS - Anismus - LCA

TPUS - Anismus - LCA

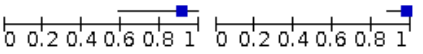
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Brusciano 2007	22	1	2	16	0.92 [0.73, 0.99]	0.94 [0.71, 1.00]		
Hainsworth 2016	87	31	4	201	0.96 [0.89, 0.99]	0.87 [0.82, 0.91]		
Martellucci 2011	3	4	1	46	0.75 [0.19, 0.99]	0.92 [0.81, 0.98]		
Ron 2012	20	7	2	73	0.91 [0.71, 0.99]	0.91 [0.83, 0.96]		
Van Gruting 2017	8	8	1	114	0.89 [0.52, 1.00]	0.93 [0.87, 0.97]		



Test 15. TPUS - PFD - LCA

TPUS - PFD - LCA

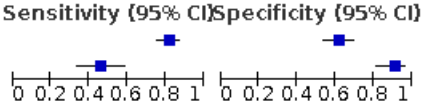
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Martellucci 2011	10	1	1	42	0.91 [0.59, 1.00]	0.98 [0.88, 1.00]		



Test 16. EVUS - Rectocele - LCA

EVUS - Rectocele - LCA

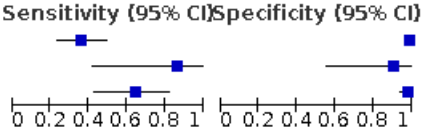
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	143	56	32	92	0.82 [0.75, 0.87]	0.62 [0.54, 0.70]		
Van Gruting 2017	30	6	35	60	0.46 [0.34, 0.59]	0.91 [0.81, 0.97]		



Test 17. EVUS - Enterocele - LCA

EVUS - Enterocele - LCA

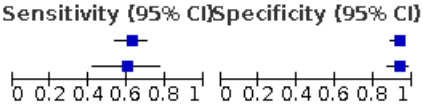
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	20	3	36	264	0.36 [0.23, 0.50]	0.99 [0.97, 1.00]		
Halligan 1996	6	1	1	9	0.86 [0.42, 1.00]	0.90 [0.55, 1.00]		
Van Gruting 2017	16	2	9	104	0.64 [0.43, 0.82]	0.98 [0.93, 1.00]		



Test 18. EVUS - Intussusception - LCA

EVUS - Intussusception - LCA

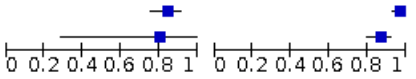
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	87	12	52	172	0.63 [0.54, 0.71]	0.93 [0.89, 0.97]		
Van Gruting 2017	20	6	13	92	0.61 [0.42, 0.77]	0.94 [0.87, 0.98]		



Test 19. EVUS - Anismus - LCA

EVUS - Anismus - LCA

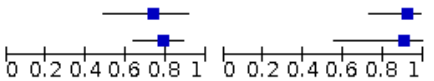
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Hainsworth 2016	77	8	14	224	0.85 [0.76, 0.91]	0.97 [0.93, 0.98]		
Van Gruting 2017	4	17	1	109	0.80 [0.28, 0.99]	0.87 [0.79, 0.92]		



Test 20. DAE - Rectocele - LCA

DAE - Rectocele - LCA

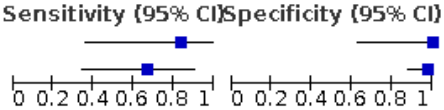
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	14	2	5	22	0.74 [0.49, 0.91]	0.92 [0.73, 0.99]		
Vitton 2011	36	1	10	9	0.78 [0.64, 0.89]	0.90 [0.55, 1.00]		



Test 21. DAE - Enterocele - LCA

DAE - Enterocele - LCA

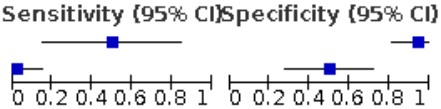
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Karaus 2000	5	0	1	8	0.83 [0.36, 1.00]	1.00 [0.63, 1.00]		
Vitton 2011	8	1	4	43	0.67 [0.35, 0.90]	0.98 [0.88, 1.00]		



Test 22. DAE - Intussusception - LCA

DAE - Intussusception - LCA

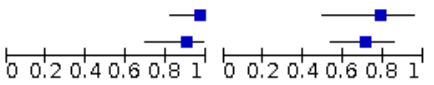
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	4	2	4	33	0.50 [0.16, 0.84]	0.94 [0.81, 0.99]		
Vitton 2011	1	11	33	11	0.03 [0.00, 0.15]	0.50 [0.28, 0.72]		



Test 23. DAE - PFD - LCA

DAE - PFD - LCA

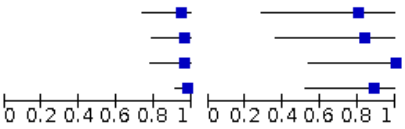
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Barthet 2000	28	3	1	11	0.97 [0.82, 1.00]	0.79 [0.49, 0.95]		
Vitton 2011	19	10	2	25	0.90 [0.70, 0.99]	0.71 [0.54, 0.85]		



Test 24. EDF - Rectocele - LCA

EDF - Rectocele - LCA

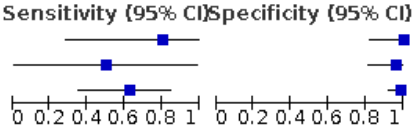
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	18	1	1	4	0.95 [0.74, 1.00]	0.80 [0.28, 0.99]		
Murad-Regadas 2008	23	1	1	5	0.96 [0.79, 1.00]	0.83 [0.36, 1.00]		
Murad-Regadas 2011	22	0	1	6	0.96 [0.78, 1.00]	1.00 [0.54, 1.00]		
Regadas 2011	75	1	2	8	0.97 [0.91, 1.00]	0.89 [0.52, 1.00]		



Test 25. EDF - Enterocele - LCA

EDF - Enterocele - LCA

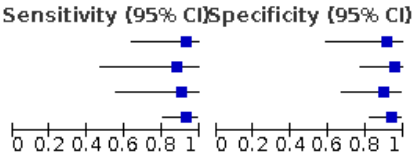
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	4	0	1	19	0.80 [0.28, 0.99]	1.00 [0.82, 1.00]		
Murad-Regadas 2011	1	1	1	26	0.50 [0.01, 0.99]	0.96 [0.81, 1.00]		
Regadas 2011	10	1	6	69	0.63 [0.35, 0.85]	0.99 [0.92, 1.00]		



Test 26. EDF - Intussusception - LCA

EDF - Intussusception - LCA

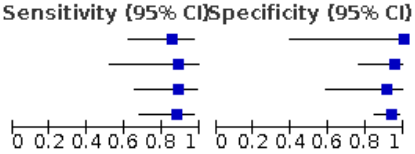
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	12	1	1	10	0.92 [0.64, 1.00]	0.91 [0.59, 1.00]		
Murad-Regadas 2008	7	1	1	21	0.88 [0.47, 1.00]	0.95 [0.77, 1.00]		
Murad-Regadas 2011	9	2	1	17	0.90 [0.55, 1.00]	0.89 [0.67, 0.99]		
Regadas 2011	38	3	3	42	0.93 [0.80, 0.98]	0.93 [0.82, 0.99]		



Test 27. EDF - Anismus - LCA

EDF - Anismus - LCA

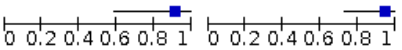
Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Miravalle 2016	17	0	3	4	0.85 [0.62, 0.97]	1.00 [0.40, 1.00]		
Murad-Regadas 2008	8	1	1	20	0.89 [0.52, 1.00]	0.95 [0.76, 1.00]		
Murad-Regadas 2011	16	1	2	10	0.89 [0.65, 0.99]	0.91 [0.59, 1.00]		
Regadas 2011	21	4	3	58	0.88 [0.68, 0.97]	0.94 [0.84, 0.98]		



Test 28. EDF - PFD - LCA

EDF - PFD - LCA

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)
Murad-Regadas 2011	10	1	1	17	0.91 [0.59, 1.00]	0.94 [0.73, 1.00]		



ADDITIONAL TABLES

Table 1. Methods of performance of imaging modalities

Test name as presented in the review	Methods of performing imaging	Alternative names as presented in the included studies
Evacuation proctography (EP)	Specific preparation is required before EP can be performed; 1 hour before the examination the participant is given 300 ml of oral liquid barium contrast for small bowel opacification and oral gastrografin® to speed up the transit time. The participant is asked to empty the bladder before the examination. The rectum is emptied with glycerine suppositories and it is filled again with 80 - 120 ml of thick barium contrast paste (potato powder and barium sulphate), which is injected through a syringe in the rectum with the participant in left lateral position. Sometimes the vagina or bladder or both are also opacified. The radiological examination starts with the participant in the upright position sitting on a special commode. A series of X-ray images or cineloops of the rectum and the anal canal are recorded at rest, on contraction, during straining and during evacuation of the contrast (Ma Kelvin 1992; Mahieu 1984; Mellgren 1994a; Shorvon 1989; Stoker 20011). When the anterior compartment needs to be visualised with EP, this could be done with the administration of hydro-soluble contrast medium in the bladder besides the barium in the rectum and vagina. This imaging technique is called cysto-colpo-defaecography (CCD), cysto-colpo-proctography (CCP) or colpo-cysto-rectography (CCRG). The technique where the bladder, vagina, small bowel and rectum are all opacified is called entero-colpo-cystodefecography (ECCP)	Dynamic evacuation proctography (DEP) Videoproctography (VP) Videodefaecography (VD) Barium proctography (BaP) Defaecation proctography Defaecography Conventional defaecography (CD) Radiographic defaecography Fluoroscopic X-ray defaecography Dynamic fluoroscopy Colpo-cysto-defaecography (CCD) Colpo-cysto-proctography (CCP) Colpo-cysto-rectography (CCRG/CCR) Entero-colpo-defecography (ECD) Entero-colpo-cysto-defaecography (ECCP)
Magnetic resonance imaging (MRI)	1. Dynamic MRI The participant is asked to have a comfortably full bladder prior to the examination. In most protocols the rectum is filled with contrast (ultrasound gel). Sometimes contrast is placed in the vagina. No contrast is used for the small bowel. For a MR-colpo-cysto-rectography the bladder is also filled with an isotone saline solution. The participant is positioned prone or supine with a	MR-defaecography MR-proctography

Table 1. Methods of performance of imaging modalities *(Continued)*

	body-phased-array receiver coil. The participant is asked to perform the rest-squeeze-relaxation-straining manoeuvre and most protocols us an evacuation phase. During the participant's attempt of rectal evacuation, whilst lying in the scanning machine with a 1 - 1.5 T closed magnet, a T2-weighted dynamic series of images is obtained in the mid-sagittal plane with 1 image every 1.2 - 2 seconds with a slide thickness of 5 mm (Colaiacomo 2009; Lienemann 1997; Piloni 2013; Pizzoferrato 2014; Stoker 2001). 2. Open-magnet MR-defaecography The participant is instructed to empty the bladder and rectum prior to the procedure. For the first series no rectal contrast is used. Imaging is performed in a 0.5 T open configuration MR system in the erect sitting position using an MR-compatible commode placed between the magnets. A flexible surface transmit-receive coil is placed under the commode upon which the participant is seated. Images are acquired in the mid-sagittal or coronal planes through the mid rectum using T1- or T2-weighted sequences during rest, squeeze and straining manoeuvres. With the participant in the left-lateral position, synthetic stool is instilled into the rectum (mashed potato starch mixed with 1% gadolinium-DTPA). The volume is inserted until the feeling of a sustained desire to defaecate is attained or to a maximum of 240 ml. The participant is reseated and images are taken during rest, squeeze and straining, as well as during evacuation (Bertschinger 2002; Dvorkin 2004; Fielding 1998; Roos 2002)	
Transperineal Ultrasound (TPUS)	For this investigation the participant is in a supine or left-lateral position with the knees semi-flexed, the legs abducted. The participant is asked to empty the bladder prior the examination. No special gynaecological chair, contrast filling or other participant preparation are required. After applying a probe cover the curved array (or convex) abdominal transducer is placed vertically on the perineum or between the labia majora, between the mons pubis and the anal margin. In the mid-sagittal plane the pubic bone, bladder, urethra, vagina, anal canal and rectum are visualised between the posterior surface of the symphysis pubis (bony landmark) and the posterior part of the levator ani. Images are required at rest, on maximal pelvic floor contraction and during maximal Valsalva manoeuvre (Dietz 2005a; Dietz 2012; Dietz 2014; Santoro 2011; Wieczorek 2011 4)	Introital ultrasound Translabial ultrasound Perineal ultrasound
Endovaginal Ultrasound (EVUS)	For this investigation the participant is in a supine position with the knees semi-flexed, the legs abducted, with the feet slightly apart from each other. No vaginal or rectal contrast needs to be used. A linear or biplane rotational transducer is inserted into the vagina in a neutral position facing the posterior compartment. 2D images and cineloops are acquired at rest and during straining (Santoro 2011; Shobeiri 2012; Wieczorek 2011 2).	
Echodefaecography (EDF)	The participant is examined in the left-lateral position after application of rectal enema. With the 360 ° rotational transducer positioned in the rectum at 6 to 7 cms from the anal verge an automatic 360 ° 3D scan is obtained in 55 seconds with a proximal-to-distal distance of 6 cms. Images are required by performing 3 scans: scan 1 at rest position without gel, scan 2 at rest–straining–at rest without gel, and scan 3 with 120 mL to 180 mL of ultrasound gel inserted into the rectum and the rest–strain–rest sequence (Murad-Regadas 2008; Murad-Regadas 2011; Regadas 2011)	Dynamic 3D anorectal ultrasonography
Dynamic anorectal endosonography (DAE)	The participant is examined in the left-lateral position. The tip of the rigid biplane transrectal probe is covered with a water-filled balloon for maintenance of the acoustic window for the ultrasound waves. The rectum is filled with 50 mL of water before the defaecation effort. The probe is slowly and manually rotated through 360 ° to identify various layers including the anal wall (mucosa, internal and external anal sphincter), the rectal wall, and the perirectal tissues (puborectalis muscle, bladder and vagina). After the initial examination, the participant is asked to make a defaecation effort while anal ul-	-

Table 1. Methods of performance of imaging modalities *(Continued)*

	trasonography is continued, leaving the probe in the same position (Barthet 2000; Vitton 2011)

Table 2. Classifications of target conditions

Target condition	Imaging technique	Classification
Rectocele	EP	(Mellgren 1994a; Kelvin 2000)
	MRI	Grade 1: rectocele depth < 2 cm (small)
		Grade 2: rectocele depth between 2 - 4 cm (moderate)
		Grade 3: rectocele depth > 4 cm (large)
	EP	(Yoshioka 1991)
		Grade 1: < 3 cm
		Grade 2: > 3 cm
	EDF	(Regadas 2011)
		Grade 1: rectocele depth < 6 mm
		Grade 2: rectocele depth between 6 - 13 mm
		Grade 3: rectocele depth > 13 mm
Enterocele	EP	(Stoker 2000)
	MRI	Grade 1: enterocele into distal half of the vagina
	TPUS	Grade 2: enterocele reaches to the perineum
		Grade 3: enterocele protruding from vagina
	EP	(Martellucci 2011)
	TPUS	Grade 1: distal part descended into the upper third of the vagina
		Grade 2: distal part descended into the middle third of the vagina
		Grade 3: distal part descended into the lower third of the vagina
	EP	(Morandi 2010)
		Grade 1: bowel extends from 2 to 4 cm below the vaginal apex (small)
		Grade 2: extension reaches 4 – 6 cm (moderate)
		Grade 3: distance is greater than 6 cm (large)
	EP	(Kelvin 1999)
		Grade 1: extension between 3 and 6 cm below the vaginal apex (small)
		Grade 2: extension between 6 and 9 cm (moderate)
		Grade 3: extension was more than 9 cm (large)

Table 2. Classifications of target conditions *(Continued)*

	MRI	(Kelvin 2000) Grade 1: extend < 3 cms below the PCL (small) Grade 2: extend from 3 to 6 cms below the PCL (moderate) Grade 3: extend > 6 cms below the PCL (large)
Intussusception	EP	(Stoker 2000)()
	MRI	Grade 1: infolding remains entirely intra-rectal (recto-rectal prolapse) Grade 2: most distal part descends into the anal canal (recto-anal prolapse) Grade 3: leading edge is protruding out of the anal canal (external rectal prolapse)
	EP	(Beer-Gabel 2004)
	TPUS	Grade 1: minimal infolding of part of the rectal wall or circumferential infolding which remains entirely intrarectal Grade 2: the leading edge extends into the orifice of the anal canal Grade 3: when the leading edge extends intra-anally
	EP	(Shorvon 1989) Grade 1: partial infolding in the rectal wall < 3 mm in width Grade 2: circumferential infolding in the rectal wall < 3 mm in width Grade 3: partial infolding in the rectal wall > 3 mm Grade 4: circumferential infolding in the rectal wall > 3 mm (intra-rectal) Grade 5: circumferential infolding in the rectal wall > 3 mm (internal anal orifice) Grade 6: circumferential infolding in the rectal wall > 3 mm (intra-anal) Grade 7: external prolapse
	EP	(Collinson 2008) (Oxford scale) Grade 1: descends no lower than proximal limit of the rectocele Grade 2: descends into the level of the rectocele, but not onto sphincter/anal canal Grade 3: descends onto sphincter/anal canal Grade 4: descends into sphincter/anal canal Grade 5: protrudes from anus
Anismus	-	No classification available
Pelvic floor descent	EP	(Bertschinger 2002)
	MRI	Grade 1: ARJ < 3 cm below PCL (small) Grade 2: ARJ 3-6 cm below PCL (moderate) Grade 3: ARJ > 6 cm below PCL (severe)

ARJ = anorectal junction; EDF = Echodefaecography; EP = evacuation proctogram; MRI = magnetic resonance imaging; PCL = pubococcygeal line; TPUS = transperineal ultrasound; US = ultrasound

Table 3. Study and patient characteristics of studies included in meta-analysis

Study ID	Period data collection	Country	Study design	Patient recruitment	Setting	Nr Centres	Nr Participants	Mean Age	Gender	Ethnicity	Symptoms	Index test	Reference standard
Barthet 2000	1997 - 1998	France	Cross-sectional DTA	Prospective	Secondary	Single	43	51 years	Female	Unknown	Symptoms involving outlet delay 100%	DAE	EP
Beer-Gabel 2004	2003	Israel	Cross-sectional DTA	Prospective	Tertiary	Single	33	58 years	Female	White	Longstanding symptoms of ODS 100%	TPUS	EP*
Beer-Gabel 2008	2008	Israel	Cross-sectional DTA	Prospective	Tertiary	Single	62	56 years	Female	White	Longstanding symptoms of ODS100%	TPUS EP	None
Beer-Gabel 2015	2011 - 2013	Israel	Cross-sectional DTA	Retrospective	Tertiary	Single	105	55 years	Female	White	Evacuation disorders (chronic constipation 77% and faecal incontinence 23%)	TPUS	EP
Brusciano 2007	2003 - 2006	Italy	Cross-sectional study	Retrospective	Secondary	Single	92	51 years	Female 77 / Male 15	White	Symptoms of ODS 100%	TPUS EVUS EP	EP for rectocele/intussusception, none for anismus
Dellemare 1994	1990 - 1992	The Netherlands	Cross-sectional	Prospective	Tertiary	Single	33	54 years	Female	Unknown	Symptoms of anterior rectocele 100%	MRI EP	None
Faggian 2013	2008 - 2011	Italy	Cross-sectional	Retrospective	Tertiary	Single	614	57.3 years	Female	Unknown	Symptoms related to pelvic floor dynamic dysfunction	MRI EP	None
Faucheron 2014	2010 - 2012	France	Cross-sectional DTA	Prospective	Tertiary	Single	50	53 years	Female	Unknown	Symptoms of ODS 100%	MRI EP	Intra-op-

Table 3. Study and patient characteristics of studies included in meta-analysis *(Continued)*

													erative results
Fiaschetti 2013	2011 - 2012	Italy	Cross-sectional	Prospective	Tertiary	Single	49	44 years	Female	Unknown	Symptoms of chronic constipation (84%), feeling of incomplete evacuation (71%) and/or faecal incontinence (20%)	MRI EP	None
Foti 2013	2007 - 2009	Italy	Cross-sectional DTA	Prospective	Secondary	Single	19	54 years	Female 17 / Male 2	Unknown	Outlet obstruction syndrome 100%	MRI	EP
Grasso 2007	2004 - 2005	Italy	Cross-sectional DTA	Prospective	Secondary	Single	43	58 years	Female	Unknown	Faecal incontinence (16%) or obstructive defaecation (86%)	TPUS EP	None
Gufler 1999	1994 - 1995	Germany	Cross-sectional	Prospective	Secondary	Single	32	61 years	Female	White	Pelvic organ prolapse 100%	MRI	None
Gufler 2004	2000	Germany	Cross-sectional	Retrospective	Secondary	Single	7	57 years	Female	White	Urinary incontinence 100%, Pelvic organ prolapse 57%	MRI EP	None
Hainsworth 2016	2011 - 2014	United Kingdom	Cross-sectional DTA	Retrospective	Tertiary	Single	393	54 years	Female	Mixed	Defaecatory dysfunction (ODS and FI) 100%	TPUS EVUS	EP
Halligan 1996	Unknown	United Kingdom	Cross-sectional DTA	Prospective	Secondary	Single	17	53 years	Female	Unknown	Symptoms of enterocele 100%	EVUS	EP
Healy 1997	Unknown	United Kingdom	Cross-sectional	Prospective	Tertiary	Single	10	61 years	Female	Unknown	Difficulty defaecating 100%	MRI EP	None
Karaus 2000	Unknown	Germany	Cross-sectional DTA	Prospective	Tertiary	Single	17	65 years	Female	Unknown	Long-standing symptoms of anorectal obstruction 100%	DAE	EP
Kelvin 2000	1999	USA	Cross-sectional	Prospective	Tertiary	Single	10	65 years	Female	Unknown	Symptoms of prolapse and pelvic floor dysfunction 100%	MRI EP	None

Table 3. Study and patient characteristics of studies included in meta-analysis *(Continued)*

Lienemann 1997	Unknown	Germany	Cross-sectional DTA	Prospective	Tertiary	Single	44	61 years	Female	Unknown	Urinary incontinence and pelvic organ prolapse 100%	MRI EP	Clinical evaluation and intraoperative results
Lienemann 2000	Unknown	Germany	Cross-sectional DTA	Prospective	Tertiary	Single	55	61 years	Female	Unknown	Isolated or combined pelvic floor descent 100%	MRI EP	Clinical examination
Martellucci 2011	2009	Italy	Cross-sectional DTA	Prospective	Tertiary	Single	54	59 years	Female	Unknown	Symptoms of ODS 100%	TPUS EP	None
Martin 2017	2009 - 2012	Spain	Cross-sectional DTA	Prospective	Tertiary	Single	40	60 years	Female 38 / Male 2	Unknown	Symptoms of ODS 100%	MRI EP	None
Matsuoka 2000	1996-1997	USA	Cross-sectional	Prospective	Tertiary	Single	9	59 years	Female	Caucasian	Chronic constipation 100%	MRI	EP
Miravalle 2016	2010 - 2014	Argentina	Cross-sectional	Prospective	Seconary	Single	24	57 years	Female	White	Symptoms of ODS 100%	EDF EP	None
Murad-Regadas 2008	2006	Brazil	Cross-sectional DTA	Prospective	Tertiary	Single	30	48 years	Female	Unknown	Symptoms of ODS 100%	EDF EP	None
Murad-Regadas 2011	2008 - 2009	Brazil	Cross-sectional DTA	Prospective	Tertiary	Single	29	43 years	Female	Unknown	Symptoms of ODS 100%	EDF EP	None
Perniola 2008	2005 - 2007	Australia	Cross-sectional DTA	Prospective	Tertiary	Single	37	53 years	Female	Unknown	Symptoms of ODS 100%	TPUS	EP
Pilkington 2012	2008 - 2009	United Kingdom	Cross-sectional DTA	Prospective	Secondary	Single	42	59 years	Female 38 / Male 4	Unknown	Symptomatic pelvic floor disorders 100%	MRI EP	None

Table 3. Study and patient characteristics of studies included in meta-analysis *(Continued)*

Poncelet 2017	2006 - 2009	France	Cross-sectional DTA	Retrospective	Unknown	Single	50	66 years	Female	Unknown	Posterior compartment dysfunction 100%	EP MRI	Composit reference standard
Regadas 2011	2009	Brazil, Venezuela, USA	Cross-sectional DTA	Prospective	Tertiary	Multi	86	53 years	Female	Unknown	Symptoms of ODS 100%	EDF EP	None
Ron 2012	2012	Israel	Cross-sectional DTA	Retrospective	Secondary	Single	102	Unknown	Female 81 / Male 21	Caucasian	Symptoms of ODS 100%	TPUS	EP
Steensma 2010	2005 - 2007	The Netherlands	Cross-sectional DTA	Prospective	Tertiary	Single	75	59 years	Female	Unknown	Symptoms related to posterior compartment prolapse. 100%	TPUS	EP
Vanbeckevoort 1999	Unknown	Belgium	Cross-sectional	Prospective	Tertiary	Single	35	65 years	Female	Unknown	Symptoms of pelvic floor descent 100%	MRI	EP
Van Gruting 2017	2014 - 2015	UK	Cross-sectional DTA	Prospective	Tertiary	Single	131	54 years	Female	White 77%	WhiteSymptoms of ODS 100%	EP MRI TPUS EVUS	Latent class analysis
Van Iersel 2017	2010 - 2011	The Netherlands	Cross-sectional	Prospective	Secondary	Single	45	64 years	Female 39 / Male 6	Unknown	Pelvic floor dysfunction of the posterior compartment	MRI	EP
Vitton 2011	2009 - 2010	France	Cross-sectional DTA	Prospective	Tertiary	Single	56	51 years	Female	Unknown	Dyschezia for > 6 months 100%	MRI DAE	EP
Weemhoff 2013	2007 - 2008	The Netherlands	Cross-sectional DTA	Prospective	Tertiary	Single	50	59 years	Female	Unknown	Symptoms of faecal incontinence 84% or obstructed defecation 16%	TPUS	EP

Table 3. Study and patient characteristics of studies included in meta-analysis *(Continued)*

Zafar 2012	2008 - 2011	United Kingdom	Cross-sectional DTA	Retrospective	Secondary	Single	16	39 years	Female 13 / Male 3	Unknown	Symptoms of ODS 100%	MRI EP	None
Zafar 2017	2012 - 2015	United Kingdom	Cross-sectional DTA	Prospective	Secondary	Single	55	59 years	Female 53 / Male 2	Unknown	Symptomes of ODS 100%	EP MRI	None

* calculations for EP as reference standard performed as secondary analysis after establishing agreement between two imaging techniques.

Table 4. Imaging characteristics included studies comparing EP and MRI

Study ID	Type of evacuation proctography	EP - patient position	EP - rectal contrast	EP - evacuation phase	Type of MRI	Type MRI scanner	MRI - participant position	MRI - rectal contrast	MRI - evacuation phase
Dellemare 1994	Radiographic defaecography	Upright	120 ml of high density BaSO4 thickened and BaSO4 contrast medium up to capacity	Yes	Dynamic MRI	Philips 1.5 Tesla Gyroscan	Prone	None	No
Faggian 2013	Entero-colpo-defaecography (ECD)	Upright	200 cc of barium paste	Yes	Supine entero-magnetic resonance (SE-MR)	1.5T closed magnet (magnetrom symphony, Siemens Germany)	Supine	200 ml ultrasound gel	Yes
Faucheron 2014	Dynamic cysto-colpo-proctography (DCP)	Upright	Semisolid contrast material of standardised consistency composed of barium suspension mixed with starch	Yes	Functional pelvic MRI	1.5 Tesla superconductive unit and a circularly polarised (quadrature) body coil (INTERA; Philips Electronics, Koninklijke, the Netherlands)	Supine	120 ml of sonographic transmission gel	Yes
Fiaschetti 2013	Colpo-cysto-defaecography (CCD)	Upright	180 - 240 ml barium paste	Yes	Magnetic resonance defaecography (MRD)	0.25 T (G-SCAN, Esaote S.p.A., Genova, Italy)	Supine and Upright	200 ml of suspension media mixed with 1 ml paramagnetic	Yes

Table 4. Imaging characteristics included studies comparing EP and MRI *(Continued)*

								contrast media	
Foti 2013	Conventional defaecography (CD)	Upright	150 - 200 ml high density barium enema	Yes	MRI	Closed-configuration superconducting unit with a 1.5-T field strength (GESigna HDx 1.5 T, GE Medical Systems, Milwaukee, WI, USA)	Supine	150 ml of ultrasound gel	Yes
Gufler 1999	Colpo-cysto-rectography (CCR)	Upright	80 mL of a barium suspension	No	Dynamic MRI	Superconductive 1.0 T Magnetom-Expert scanner (Siemens, Erlangen, Germany)	Supine	None	No
Gufler 2004	Colpo-cysto-proctography (CCP)	Upright and supine	barium suspension	No	Dynamic MRI	1.0 T Magnetom- Expert scanner (Siemens, Erlangen, Germany)	Supine	None	No
Healy 1997	Evacuation proctography	Upright	120 ml of barium paste	Yes	Dynamic MR imaging	1.5-T superconducting magnet system (Signa: General Electric Medical Systems, Milwaukee, WI)	Supine	Soft rubber tube 5 mm in diameter	No
Kelvin 2000	Dynamic fluoroscopic cysto-colpo-proctography	Upright	200 ml of a thick barium paste	Yes	Dynamic MR cysto-colpo-proctography	1.5-T superconductive unit and a circularly polarised (quadrature) body coil (Vision; Siemens, Erlangen, Germany)	Supine	200 ml sonographic transmission gel	Yes
Lienemann 1997	Dynamic fluoroscopy	Upright	200 ml of barium; Micropaque, Guerbet	Yes	MRI	1.5-T superconductive magnet unit (Vision, Siemens, Erlangen, Germany)	Supine	200 ml sonography gel	Yes
Lienemann 2000	Dynamic cysto-procto-graphy	Upright	Barium suspension (approximately 200 ml)	Yes	MR Colpo-cysto-rectography	1.5 Tesla System TM (Vision, Siemens Corp., Erlangen, Germany)	Supine	Sonography gel (approximately 50 ml)	Yes
Martin 2017	Videodefaecography (VD)	Upright	200 gr of potato puree flakes, liquid barium sulphate and 700 ml of water	Yes	MR defaecography	Siemens Magnetom Sonata closed MRI of 1.5 Tesla (T)	Supine	100 g of potato puree flakes, 400 g of barium sulphate, 7 ml of gadolinium, and water un-	Yes

Table 4. Imaging characteristics included studies comparing EP and MRI *(Continued)*

								til a solution of 450 ml was reached	
Matsuoka 2000	Videoproctography (VP)	Upright	50 ml liquid barium and up to 100 ml a thick barium paste	Yes	Dynamic pelvic MRI (DPMRI)	1T Picker Vista Edge MRI (Picker, Highland Hights, Ohio, USA)	Prone	50 ml air	No
Pilkington 2012	Barium proctography (BaP)	Upright	Barium paste	Yes	MR proctography	1 T magnet (Phillips Intera)	Supine	Ultrasound gel	Yes
Poncelet 2017	X-ray defaecography	Upright	120 ml of barium sulfate mixed with Smecta®	Yes	MR defaecography	1.5 T Signa (GE or Phillips)	Supine	200 - 250 ml of ultrasound gel	Yes
Vanbeckevoort 1999	Dynamic colpo-cysto-defaecography (CCD)	Upright	Barium	Yes	Dynamic MRI	1.5 T system Magnetom Vision, Siemens Medical Systems, Erlangen, Germany	Supine	100 ml ultrasound gel	No
Van Gruting 2017	Evacuation proctogram (EP)	Upright	80 ml to 120 ml of barium paste	Yes	Dynamic MRI	Closed MRI scanner with a 1.5 T magnet (Siemens Avanto)	Supine	120 ml of ultrasound gel	Yes
Van Iersel 2017	Dynamic conventional (entero-colpo) defaecography (CD)	Upright	300 ml of barium paste (barium sulphate, liquid polibar) and water (35% wt/vol)	Yes	Dynamic MR defaecography (D-MRI)	1.5-T closed magnet (Intera rel.2.6.3, Philips, Best, The Netherlands)	Supine	200 ml ultrasonographic gel	Yes
Vitton 2011	Conventional defaecography	Upright	300 ml contrast	Yes	Dynamic MRI defaecography	1.5-T superconductive unit and a circularly polarised (quadrature) body coil (Intera; Philips Medical Systems, Best, The Netherlands)	Supine	100mL of sonographic transmission gel	Yes
Zafar 2012	Evacuation proctography (EP)	Upright	Yes, type and volume not described	Yes	Magnetic resonance defaecography (MRD)	1.5 Tesla closed magnet Siemens Symphony scanner	Supine	Yes, ultrasound gel, volume not described	Yes

Table 4. Imaging characteristics included studies comparing EP and MRI *(Continued)*

Zafar 2017	Evacuation proctography (EP)	Upright	Thickened barium paste	Yes	Magnetic resonance defaecography (MRD)	1.5 T closed magnet (Siemens, Germany)	Supine	Ultrasound Gel (120 ml)	Yes

Table 5. Imaging characteristics included studies comparing EP and pelvic floor ultrasound

Study ID	Type of evacuation proctography	EP - participant position	EP - rectal contrast	EP - evacuation phase	Type of pelvic floor ultrasound	Type of ultrasound scanner	Type of ultrasound probe	Ultrasound - participant position	Ultrasound - rectal contrast	Ultrasound - evacuation phase
Barthet 2000	Defaecography	Upright	Yes, type and volume not described	Yes	Dynamic anal endosonography (DAE)	PVL-625RT Toshiba	7MHz rigid linear endoanal probe	Left-lateral	n = 28 no contrast n = 15 50 ml water	No
Beer-Gabel 2004	Evacuation proctography	Upright	120 ml barium paste	Yes	Dynamic 2D transperineal ultrasound	HDI 3000, Advanced Technology Laboratories, USA	Curvilinear transducers (C 4-7 and C 8-12)	Left-lateral	50 ml ultrasonographic coupling gel	Yes
Beer-Gabel 2008	Dynamic evacuation proctography (DEP)	Upright	150 mL of contrast medium	Yes	Dynamic 2D transperineal ultrasonography (DTP-US)	Logiq 9, GE Healthcare UK	a curvilinear C4–7 or a C8–12 transducer	Left-lateral	50 mL of ultrasonographic coupling gel	Yes in some cases
Beer-Gabel 2015	Evacuation proctography	Upright	120 ml barium paste	Yes	Dynamic 2D transperineal ultrasonography (DTP-US)	BK medical, profocus	curvilinear 5 – 8 MHz probe	Left-lateral	50 mL of ultrasonographic coupling gel	No
Brusciano 2007	Defaecography or Entero-colpo-	Upright	Barium, amount not described	Yes	Dynamic perineal US	BK medical	Linear 5- to 8-mHz probe	Supine	None	No

Table 5. Imaging characteristics included studies comparing EP and pelvic floor ultrasound *(Continued)*

	defaecography									
Grasso 2007	Colpo-cysto- defaecography (CCD)	Upright	200 mL barium paste	Yes	Introital ultrasound	Sonoline Antares (Siemens AG, Erlangen, Germany)	6.2-MHz EC9-4 probe	Semi-recumbent position (110 ° sitting angle)	None	No
Halligan 1996	Evacuation proctography	Upright	120 ml of barium paste	Yes	Vaginal endosonography	BK Medical	Type 1850 rectal endoprobe	Left-lateral	None	No
Hainsworth 2016	Defaecation proctography	Upright	Paste (mixture of Baritop®, Readybrek® and warm water)	Yes	transperineal utrasound transvaginal ultrasound	BK Medical	Curved array probe (6 MHz) Linear array endoscopic probe (12 MHz)	Supine	None	No
Karaus 2000	Defaecography	Upright	200 ml of contrast medium	Yes	Anorectal endosonography (DAE)	Kontron Instruments, AI 52000S, Neufahrn, Germany	Transversal sector scanner and a sagittal curved array scanner (65 MHz ER-BI-T, 7,5 MHz ER-BI-S)	Left-lateral	None	No
Martelluc-ci 2011	Dynamic evacuation proctography (DEP)	Upright	200 ml of thick barium sulphate paste	Yes	Dynamic 2D transperineal ultrasound (DT-PU)	BK Medical	Type 8802, 6 MHz, convex probe	Supine	n = 15: ultrasonographic gel n = 39: None	No
Miravalle 2016	Defaecography	Upright	150 ml of barium paste	Yes	Echodefaecography (EDF)	BK Medical	Type 2050, endoprobe	Left-lateral	Ultrasound gel (120 ml)	Yes
Murad-Regadas 2008	Conventional defaecography (DF)	Upright	300 ml barium paste	Yes	Echodefaecography (EDF)	BK Medical	Type 2050, endoprobe	Left-lateral	Ultrasound gel (120 – 180 ml)	No

Table 5. Imaging characteristics included studies comparing EP and pelvic floor ultrasound *(Continued)*

Murad-Regadas 2011	Defaecography	Upright	200 ml barium paste	Yes	Dynamic anorectal ultrasonography (Dynamic 3-DAUS)	BK Medical	Type 2050, endoprobe	Left-lateral	Ultrasound gel (120–180 ml)	No
Perniola 2008	Defaecation proctography	Upright	Barium or Liquid Polybar Plus followed by a Liquid Polybar/starch mixture	Yes	4D Translabial ultrasound	GE Kretz Voluson 730 Expert system	4D abdominal transducer	Supine	No	No
Regadas 2011	Defaecography	Upright	Barium paste 150 mL	Yes	Echodefaecography (EDF)	BK Medical	Type 2050, endoprobe	Left-lateral	Ultrasound gel (120–180 ml)	No
Ron 2012	Evacuation proctography	Upright	200 ml barium paste	Yes	Transperineal ultrasound	Hitachi, Hi Vision	Small convex probe	Left-lateral	Ultrasound gel 120 ml	Yes
Steensma 2010	Evacuation proctography	Upright	Liquid barium contrast, amount not described	Yes	4D Transperineal ultrasound	GE Kretz Voluson 730 expert system	abdominal 4–8 MHz transducer with 3D data acquisition	Supine	None	No
Van Gruting 2017	Evacuation proctogram (EP)	Upright	80 ml to 120 ml of barium paste	Yes	2D transperineal ultrasound 2D endovaginal ultrasound	BK Medical	Type 8802, 3.5-6.0 MHz, focal range 10-135 mm Type 8838, 6–12 MHz, focal range 3 - 60 mm, contact surface 65 x 5.5 cm	Supine	None	No
Vitton 2011	Conventional defaecography	Upright	300 ml barium	Yes	Dynamic anal endosonography (DAE)	model EUP-U533; Hitachi Medical Systems, Tokyo, Japan	rigid biplane transrectal probe with a frequency of 7 MHz	Left-lateral	50 ml water	No

Table 5. Imaging characteristics included studies comparing EP and pelvic floor ultrasound *(Continued)*

Weemhoff 2013	Evacuation proctogram	Upright	Barium	Yes	2D Transperineal ultrasound	GE Kretz Voluson 730 Expert	4- to 8-mHz transabdominal curved 2D transducer	Supine	None	No

Table 6. Rectocele definition and cut-off values used in included studies

Study ID	Rectocele definition	EP - rectocele cut-off value	MRI - rectocele cut-off value	Ultrasound - rectocele cut-off value
Barthet 2000	*DAE:* Rectocele was identified if the rectal wall bulged into the vaginal lumen *Proctography:* Rectocele was defined as any anterior bulge outside the extrapolated line of the anterior rectal wall	Any	N/A	Any
Beer-Gabel 2004; Beer-Gabel 2008; Beer-Gabel 2015	A rectocele was defined as any outpouching of the anterior rectal wall occurring during evacuation or straining. Rectoceles were assigned to 1 of 2 groups based on depth. Mid-sized rectoceles were 2 – 4 cm deep and large rectoceles were defined as deeper than 4 cm.	≥ 20 mm rectocele depth	N/A	≥ 20 mm rectocele depth
Brusciano 2007	*Transperineal ultrasound:* Rectocele appeared as a semi circumferential anterior hypoechogenic area between the rectum and the vagina on straining *Proctography:* Rectocele was defined as a bulge of the anterior rectal wall more than 2 cm in size with or without entrapping of barium on straining.	≥ 20 mm rectocele depth	N/A	*Transperineal:* > 10 mm depth
Dellemare 1994	The distance between the projection of the anorectal junction and the anterior rectal wall on the baseline is defined by us as the quantitative size of the anterior rectocele. Grade 0 = absent, Grade I = moderate, Grade II = severe	Any	Any	N/A
Faucheron 2014	A rectocele was diagnosed if the anterior margin of the rectal wall bulge was more than 3 cm anterior to a line drawn along the long axis of the anterior anal canal	> 30 mm rectocele depth	> 30 mm rectocele depth	N/A
Fiaschetti 2013	The rectocele, either anterior or posterior, was evaluated by drawing a line parallel to the anterior or posterior wall of the anal canal and measuring the distance between this line and the widest point of bulging	Any	Any	N/A
Foti 2013	No definition provided. Grade 1: < 2 cm depth, Grade 2: 2 - 4 cm depth, Grade 3: > 4 cm depth	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth	N/A
Grasso 2007	*Introital ultrasound:* Rectocele was identified on sagittal scans as a bulging of the hypoechoic anterior rectal wall, detectable at rest and/or more evident during the straining manoeuvre	> 0 mm depth	N/A	> 0 mm depth

Table 6. Rectocele definition and cut-off values used in included studies *(Continued)*

	Proctography: A rectocele was defined as a rectal bulge, based on the maximum depth reached by a line traced at 90 ° with respect to a tangential line traced along the anterior wall of the anal canal			
Gufler 1999 **Gufler 2004**	Outpouching of the rectal wall, usually ventrally	≥ 10 mm rectocele depth	≥ 10 mm rectocele depth	N/A
Hainsworth 2016	*Proctography:* A bulge of the anterior rectal wall beyond the projected anterior rectal wall *Transperineal:* Bulging of the anterior rectal wall during the Valsalva manoeuvre *Transvaginal:* Protrusion of the anterior rectal wall over the perineal body	≥ 20 mm rectocele depth	N/A	*Transperineal:* ≥ 20 mm rectocele depth *Transvaginal:* Protrusion of the anterior rectal wall over the perineal body
Healy 1997	A rectocele was defined as an outpouching of the anterior rectal wall occurring during evacuation or straining was identified and, if present, its depth was measured perpendicular to the expected position of the anterior rectal wall.	Any	Any	N/A
Kelvin 2000	A rectocele was defined as any rectal protrusion anterior to a line extended upward through the anal canal. Rectoceles were graded as small if they measured < 2 cm in extent, moderate if they measured from 2 to 4 cm in extent, and large if they measured 4 cm or more in extent	Any	Any	N/A
Lienemann 1997	A bulge of more than 3 cm measured as the distance between the extended line of the anterior border of the anal canal and the tip of the rectocele was interpreted as a rectocele	> 30 mm rectocele depth	> 30 mm rectocele depth	N/A
Martellucci 2011	*Ultrasound:* A rectocele was defined as a discontinuity in the anterior anorectal muscularis, resulting in a herniation of rectal contents into the vagina. Rectocele depth was measured perpendicular to a line projected along the expected contour of the anterior rectal wall *Proctography:* A rectocele was diagnosed when the anterior rectal and posterior vaginal wall herniated into the lumen of the vagina. Its depth was assessed by the length of the segment drawn from this axis to the point of maximum convexity of the rectocele. A rectocele was considered small (first degree) if it was < 2 cm in depth, moderate (second degree) if it was 2 – 4	≥10 mm rectocele depth	N/A	≥10 mm rectocele depth

Table 6. Rectocele definition and cut-off values used in included studies *(Continued)*

	cm in depth and large (third degree) if it was more than 4 cm in depth			
Martin 2017	Bulge extending beyond the expected line of the anterior rectal wall (grade I: < 2 cm, grade II: 2 – 4 cm, grade III: > 4 cm)	Any	Any	N/A
Matsuoka 2000	Protrusion beyond the extrapolated rectal wall	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth	N/A
Murad-Regadas 2008; **Murad-Regadas 2011** **Regadas 2011**	*EDF:* Displacement of the lower rectal wall and bulging into the vaginal lumen at the point of maximal straining. *Proctography:* Anorectocele was defined as any outpouching of the anterior upper anal canal and rectal wall occurring during straining	> 0 mm depth	N/A	> 0 mm depth
Perniola 2008	Rectocele was defined as a discontinuity in the anterior anorectal muscularis, resulting in a herniation of rectal contents into the vagina	> 0 mm depth	N/A	> 10 mm depth
Pilkington 2012	Rectocele size was measured as the maximum length from an extended anterior wall of the anal canal	Any	Any	N/A
Poncelet 2017	Rectocele was considered present when the anterior bulge of the rectal wall was larger than 25 mm and further considered as grade 1 when between 25 and 50 mm and grade 2 if larger than 50 mm	> 25 mm	> 25 mm	N/A
Steensma 2010	*Ultrasound:* Rectocele was defined as a defect in the rectovaginal septum. This was seen as a sharp discontinuity in the ventral contour of the anorectal muscularis, which resulted in a herniation *Procography:* Rectocele was defined as a herniation of the anterior rectal wall into the lumen of the vagina. Rectocele depth was measured perpendicular to a line projected along the expected contour of the anterior rectal wall	> 0 mm depth	N/A	≥10 mm rectocele depth
Vanbeckevoort 1999	Rectocele was defined as an outpouching of the anterior wall more than 3 cm during straining down	> 30 mm rectocele depth	> 30 mm rectocele depth	N/A
Van Gruting 2017	A rectocele is a bulging of the anterior rectal wall into the vagina. Its depth was measured as the maximum depth perpendicular to the expected contour of the anterior rectal wall	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth	≥10 mm rectocele depth
Van Iersel 2017	A rectocele was defined as a protrusion during evacuation or during maximal straining of the rectal wall of more than 20 mm anterior to a longitudinal line parallel to the axis of the anal canal	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth	N/A

Table 6. Rectocele definition and cut-off values used in included studies *(Continued)*

Vitton 2011	A rectocele was diagnosed if the ventral rectal wall bulged by more than 2 cm into the vaginal lumen during straining to defaecate.	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth	≥ 20 mm rectocele depth
Weemhoff 2013	Rectocele was defined as bulging of the anterior rectal wall into the vagina. The maximum depth of the bulging rectocele was measured perpendicular to the expected contour of the anterior rectal wall. Grade 1 was a rectocele with a depth below 2 cm. Grade 2 rectocele had a depth between 2 and 4 cm. In a grade 3 rectocele, the depth of the bulge exceeded 4 cm	≥ 20 mm rectocele depth	N/A	≥ 20 mm rectocele depth
Zafar 2017	A rectocele was defined as a protrusion of the anterior rectal wall into the lumen of the vagina. These were categorised according to size (small (< 2 cm), medium (2 – 4 cm) and large (> 4 cm)) and whether or not they retained contrast	> 20 mm rectocele depth	> 20 mm rectocele depth	N/A

N/A = not applicable

Table 7. Enterocele definitions and cut-off value used in included studies

Study ID	Enterocele definition	EP - Enterocele cut-off value	MRI - Enterocele cut-off value	Ultrasound - Enterocele cut-off
Beer-Gabel 2004; **Beer-Gabel 2008**; **Beer-Gabel 2015**	Enteroceles were readily identified as small bowel loops visible in the region of the rectovaginal septum. Peritoneoceles were defined as an enlarged rectovaginal septum without visible small-bowel loops being present	A cul-de-sac hernia was present when there was prolapse of the posterior vaginal wall (or of the vaginal vault) during straining	N/A	A cul-de-sac hernia was present when there was prolapse of the posterior vaginal wall (or of the vaginal vault) during straining
Brusciano 2007	*Transperineal Ultrasound:* Enterocele appeared as an oval hypoechogenic area (intestinal fluid) surrounded by an hyperechogenic layer (intestinal wall) between the anorectum and the vagina, which was more evident on straining *Proctography:* Enterocele/sigmoidocele was defined as a bowel loop descending in an enlarged recto-vaginal space	Bowel loops descending in an enlarged recto-vaginal space	N/A	Bowel loops between the anorectum and the vagina
Faggian 2013	Enterocele is a descent of the small bowel loops, the peritoneal fat or the sigmoid colon into the rectogenital space above the superior portion of the vaginal dome	Small bowel loops, the peritoneal fat or the sigmoid colon into the rectogenital space	Small bowel loops, the peritoneal fat or the sigmoid colon into the rectogenital space	N/A
Faucheron 2014	Enlargement of the rectovaginal septum indicates descent of small bowel, sigmoid or great omentum between the rectum and the vagina	Enlargement of the recto-vaginal septum	Small bowel loops below PCL	N/A

Table 7. Enterocele definitions and cut-off value used in included studies *(Continued)*

Fiaschetti 2013	Enterocele was defined as small bowel loops below the PCL	Bowel loops > 10 mm below the PCL	Bowel loops > 10 mm below the PCL	N/A
Foti 2013	No definition provided. Grade 1: < 3 cm, Grade 2: 3 - 6 cm, Grade 3: > 6 cm below PCL	Small bowel or recto-sigmoid below PCL	Small bowel or recto-sigmoid below PCL	N/A
Gufler 1999Gufler 2004	Descending of bowel loops below the PCL	Small bowel loops below PCL	Small bowel loops below PCL	N/A
Hainsworth 2016	*Proctography:* The descent of contrast-filled small bowel loops onto the rectum to touch the rectal wall *Transperineal:* The descent of bowel to fill the rectovaginal space. It is graded according to the extent of descent (grade I - most distal part descended into the upper third of the vagina; grade II - middle third or grade III – the lower third) *Transvaginal:* Presence of bowel between the rectum and vaginal wall.	The descent of contrast-filled small bowel loops onto the rectum to touch the rectal wall	N/A	*Transperineal:* most distal part at least into the upper third of the vagina *Transvaginal:* Presence of bowel between the rectum and vaginal wall
Halligan 1996	*Ultrasound:* A diagnosis of enterocele was made if the rectum became obscured by bowel loops during straining *Proctography:* A diagnosis of enterocele was made if the vaginal marker was displaced away from the anterior rectal wall during evacuation	Displacement of vaginal marker during evacuation	N/A	Rectum obscured by bowel loops during straining
Kelvin 2000	An enterocele or sigmoidocele was defined as descent of the small bowel or sigmoid colon below the pubococcygeal line. Enteroceles or sigmoidoceles were graded as small if they extended less than 3 cm below the pubococcygeal line, moderate if they extended from 3 to 6 cm below this line, and large if they extended 6 cm or more below this line. A peritoneocele was defined as herniation of the peritoneal cul-de-sac with or without contained small bowel or sigmoid colon, and was measured in the same manner as enterocele and sigmoidocele	Descent of the small bowel or sigmoid colon below the PCL	Descent of the small bowel or sigmoid colon below PCL	N/A
Lienemann 1997	*Proctography:* An enlarged space between the vagina and the anterior wall of the rectum *MRI*	Enlarged space between the vagina and the anterior wall of the rectum	Widening of the rectovaginal space or deepening of the pouch of Douglas below PCL	N/A

Table 7. Enterocele definitions and cut-off value used in included studies *(Continued)*

	Widening of the rectovaginal space or deepening of the pouch of Douglas with or without small bowel loops beyond the reference line (PCL)			
Lienemann 2000	*MRI* A descent of parts of the peritoneum below the pubococcygeal reference line was diagnosed as being an enterocele. *Proctography:* An increase in the distance between the delineated vagina and rectum during straining compared to relaxation.This expansion should extend below the pubococcygeal reference line and show a sagittal diameter of more than 2 cm	Small bowel below PCL	> 2 cm width of the rectovaginal space below the PCL	N/A
Martellucci 2011	*Ultrasound:* Enterocele was diagnosed by the descent of intra-abdominal contents on Valsalva manoeuvre. It was defined as Grade 1 when the most distal part descended into the upper third of the vagina, Grade 2 when the distal part descended into the middle third and Grade 3 when the distal part descended into the lower third *Proctography:* An enterocele was diagnosed when the contrast-filled small bowel loops descend between the rectum and vagina	Small bowel loops descending between the rectum and vagina	N/A	Intra-abdominal contents into at least the upper third of the vagina
Martin 2017	*MRI:* Pelvic herniation during defaecation formed by an abnormally deep Douglas pouch contained by the small bowel, sigmoid colon or peritoneal fluid / mesenteric fat *Proctography:* Small bowel or sigmoid filling an abnormal peritoneal space in the pelvic floor	Small bowel or sigmoid colon which descend into an abnormal peritoneal cavity during defaecation	Pelvic herniation during defaecation formed by an abnormally deep Douglas pouch	N/A
Murad-Regadas 2011; **Regadas 2011**	*EDF:* Enterocele was recognised when the small bowel was positioned below the pubococcygeal line *Defaecography:* Enterocele and sigmoidocele were diagnosed as herniations of the peritoneum (containing the small bowel or sigmoid colon) into the pelvis. Extension of the loop of the small bowel or sigmoid below the ischiococcygeal line was considered significant	Small bowel or sigmoid below the ischiococcygeal line	N/A	Small bowel below the pubococcygeal line
Pilkington 2012	Descent of small bowel to perineum or below projected peritoneal reflection	Any	Any	N/A

Table 7. Enterocele definitions and cut-off value used in included studies *(Continued)*

Poncelet 2017	Grade 1 reaches the upper part of the vagina, grade 2 in case of involvement of the lower part of the vagina up to the vulva, and grade 3 in case of externalisation	Upper part of the vagina	Upper part of the vagina	N/A
Steensma 2010	Enterocele was described as a herniation of small bowel or rectosigmoid into the vagina. Grade 1 Most distal part descending into upper ⅓ of the vagina. Grade 2 Most distal part descending into middle ⅓ of the vagina. Grade 3 Most distal part descending into lower 1 ⅓ of the vagina	Herniation of small bowel or rectosigmoid into the vagina	N/A	Abdominal contents developed anterior to the anorectal junction and extended into the vagina
Vanbeckevoort 1999	Enterocele was diagnosed when bowel loops or peritoneal fat were interposed between the upper vagina and adjacent rectum, or when bowel loops prolapsed below the pubococcygeal line	Bowel loops or peritoneal fat below PCL	Bowel loops or peritoneal fat below PCL	N/A
Van Gruting 2017	An enterocele is the descent of small bowel loops or rectosigmoid between the rectum and vagina on Valsalva manoeuvre	Small bowel or rectosigmoid below PCL	Small bowel or rectosigmoid below PCL	*Transperineal ultrasound:* Small bowel or rectosigmoid visible below the superior inferior border of the symphysis pubis on Valsalva *Endovaginal ultrasound:* Small bowel loops or sigmoid colon visible in the region of the rectovaginal septum
Van Iersel 2017	Small bowel in the recto-vaginal septum extending below the PCL	Small bowel below PCL	Small bowel below PCL	N/A
Vitton 2011	An enterocele was defined as an internal herniation of the peritoneal sac into the rectovaginal space below the pubococcygeal line	Peritoneal sac the in the rectovaginal space below the PCL	Small bowel below the PCL	Peritoneal sac in the rectovaginal space
Weemhoff 2013	Enterocele was defined ad a herniation of the peritoneal cavity with abdominal contents between the rectum and vagina. Grade 1 enterocele extended into the upper half of the vagina. Grade 2 enterocele reached the perineum, and grade 3 enterocele protruded out of the vagina	Abdominal content between rectum and vagina	N/A	Abdominal content between rectum and vagina

N/A = not applicable; PCL = Pubococcygeal line

Table 8. Intussusception definition and cut-off value used in included studies

Study ID	Intussusception definition	EP - Intussusception cut-off value	MRI - Intussusception cut-off value	Ultrasound - Intussusception cut-off value
Barthet 2000	Rectal intussusception was defined as a circumferential descent of the entire thickness of the rectal wall without passing through the anal canal	Full thickness circumferential descent of rectal wall	N/A	Full thickness circumferential descent of rectal wall
Beer-Gabel 2004; Beer-Gabel 2015	Intussusception was defined as folding of the rectum into either the rectum (recto-rectal) or in contact with the anus (recto-anal) or penetrating to the anal canal (intra-anal)	Any	N/A	Any
Brusciano 2007	*Transperineal ultrasound:* Recto-rectal and recto-anal intussusception are detectable as a hyperechoic mass, i.e. the prolapsed rectal mucosa, during forcible straining commencing at the level of puborectalis sling, or just above, and possibly surrounded by 1 or 2 hypoechoic rings represented by the intussuscepted muscularis propria of the rectum *Proctography:* Recto-rectal and rectoanal (i.e. internal mucosal prolapse) intussusception was defined as intraluminal folding at the level of the rectum and/or in the anal canal, respectively	Circumferential or semilunar	N/A	Circumferential or semilunar
Faucheron 2014	Not described	Any	Any	N/A
Fiaschetti 2013	Invagination of the rectal wall	Full thickness or mucosal only	Full thickness or mucosal only	N/A
Foti 2013	Intussusception is defined as internal intra-rectal prolapse if the invagination is limited to the rectum or as internal intra-anal prolapse if the apex enters the anal canal and remains inside it during straining. External rectal prolapse is invagination of the rectal wall through the anal canal	Intra-rectal or intra-anal invagination	Intra-rectal or intra-anal invagination	N/A
Grasso 2007	*Introital ultrasound:* Intussusception was diagnosed on sagittal scans if there was an infolding of the hypoechoic anterior and posterior rectal walls during straining *Proctography:* Intussusception is an internal prolapse and may be intrarectal or intra-anal	Any	N/A	Any
Hainsworth 2016	*Proctography:* Intussusception was graded according to the Oxford Radiological Classification (Grades I and II, recto-rectal (normal); Grades III and IV, recto-anal (pathological); Grade V, rectal prolapse)	Grades III and IV, recto-anal (pathological)	N/A	Grade III – IV: infolding reached beyond the inferior edge of puborectalis but stopped be-

Table 8. Intussusception definition and cut-off value used in included studies *(Continued)*

	Transvaginal: Grade I – II: infolding rectal wall ceased prior to the inferior edge of the puborectalis; Grade III – IV: infolding reached beyond the inferior edge of puborectalis but stopped before the perineal body; Grade V: infolding rectal wall protruded beyond the perineal body			fore the perineal body
Martellucci 2011	*Ultrasound:* Intussusception and rectal prolapse were diagnosed if there was an infolding of the hypoechoic anterior or posterior rectal wall during straining. It was classified as recto-rectal, ano-rectal and external *Proctogram:* An intussusception was defined as a circumferential intraluminal folding of the rectal wall above the anal canal(recto-rectal prolapse), involving the anal canal (rectoanal prolapse) or coming out through the anal verge (external rectal prolapse).	Circumferential intraluminal folding of the rectal wall	N/A	Anterior or posterior infolding
Martin 2017	*MRI:* Descending full-thickness invagination of the rectal wall insufficient in descent to appear beyond the anal verge as an external rectal prolapse *Proctography:* Rectum showing a funnel-shaped depression within the anal canal during push	Funnel-shaped depression within the anal canal during push	Full-thickness invagination of the rectal wall	N/A
Matsuoka 2000	Circumferential infolding during evacuation or pushing	Circumferential infolding	Circumferential infolding	N/A
Murad-Regadas 2008; **Murad-Regadas 2011**; **Regadas 2011**	*EDF:* Intussusception was clearly identified on echodefaecography by observing the rectal wall layers protruding through the rectal lumen *Defaecography:* Intussusception was defined as invagination of the rectal wall occurring during straining but not passing through the anal canal.	Any	N/A	Any
Perniola 2008	Full-thickness invagination of the rectal wall into the anal canal	Full-thickness invagination	N/A	Full-thickness invagination
Pilkington 2012	Full-thickness circumferential invagination of the rectal wall into the rectum or anal canal.	Full-thickness circumferential invagination	Full-thickness circumferential invagination	N/A
Poncelet 2017	Rectal prolapse was categorised intrarectal (grade 1), rectoanal (grade 2) or external (grade 3)	Any	Any	N/A

Table 8. Intussusception definition and cut-off value used in included studies *(Continued)*

Steensma 2010	Intussusception was defined as an infolding of the rectal wall into the rectum or anus. When an external component was present, it was called complete rectal prolapse. Grade 1 Most distal part remains completely intrarectal. Grade 2 Most distal part descending into anal canal	Intra-rectal or intra-anal infolding	N/A	Intra-rectal or intra-anal infolding
Van Gruting 2017	An intussusception is an intraluminal folding of the rectal wall. An intussusception was defined as a full thickness circumferential infolding that either extended in to the rectum (grade 1), anal canal (grade 2) or externally (grade 3).	Full thickness circumferential infolding of the rectal wall during straining	Full thickness circumferential infolding of the rectal wall during straining	Full thickness circumferential infolding of the rectal wall during straining
Van Iersel 2017	As a circumferential rectal wall invagination or infolding descending toward the anal canal	Circumferential invagination	circumferential invagination	N/A
Vitton 2011	Rectal intussusception was defined as a circumferential descent of the entire thickness of the rectal wall without passing through the anal canal	Full-thickness circumferential infolding	Mucosal or full thickness	Mucosal or full thickness
Weemhoff 2013	Intussusception was defined as invagination of the proximal rectal wall during defaecation	Present	N/A	Present
Zafar 2017	Rectal intussusception was defined as circumferential descent of the entire thickness of the rectal wall or mucosa, which might extend into the anal canal but not through the anal verge	Circumfirential descent (full-thickness or mucosa)	Circumfirential descent (full-thickness or mucosa)	N/A

N/A = not applicable

Table 9. Anismus definition and cut-off value used in included studies

Study ID	Animus definition	EP - Anismus cut-off value	MRI - Anismus cut-off value	Ultrasound - Anismus cut-off value
Brusciano 2007	*Transperineal ultrasound:* The relaxation of the puborectalis muscle was defined as the difference (Δ) in millimetres between the distance of the inner edge of the puborectalis muscle posteriorly and the probe at rest (R) and on straining (S) [Δ = R−S] *Proctography:* Anismus was defined as either a lack of shortening and widening of the anal canal on straining, due to lack of puborectalis muscle relaxation, or to a paradoxical contraction of the muscle itself, detectable as indentation over the anorectal channel, mimicking an endoluminal defect shown by the barium paste	Lack of shortening and widening of the anal canal on straining	N/A	If Δ was ≤ 1 mm
Foti 2013	In the case of spastic pelvic floor syndrome (anismus, pelvic floor dyssynergia), during evacuation, the ARA tends to become more acute rather than	ARA more acute during straining	ARA more acute during straining	N/A

Table 9. Anismus definition and cut-off value used in included studies *(Continued)*

	obtuse, which indicates a failed release of the puborectal muscle			
Grasso 2007	*Ultrasound:* Dyssynergia was defined as a failure to open the ARA during straining; we considered the puborectalis muscle dyssynergic when the straining/rest ratio, calculated on the ARA, was ≤ 1, as the ARA is normally higher during straining than it is during squeezing *Proctography:* Impaired evacuation during proctography is highly specific for the diagnosis of anismus; we defined impaired evacuation as the inability to evacuate ⅔ of the contrast medium within 30 seconds	Inability to evacuate ⅔ of the contrast medium within 30 seconds	N/A	Straining/rest ratio ≤ 1
Healy 1997	The anorectal angle, defined as the angle formed by the longitudinal axis of the anal canal and the posterior border of the rectal wall, was measured at rest and during evacuation or maximal straining	ARA more acute during straining	ARA more acute during straining	N/A
Hainsworth 2016	*Proctography:* Failure to relax or a paradoxical increase in the anorectal angle during attempted evacuation *Transperineal/transvaginal:* Failure to relax or a paradoxical increase in the anorectal angle on bearing down	Failure to relax or a paradoxical increase in the ARA on evacuation	N/A	Failure to relax or a paradoxical increase in the ARA on bearing down
Martellucci 2011	*Ultrasound:* The ARA was measured at the intersection of a line forming the longitudinal axis of the anal canal with that of the posterior border of the rectal wall. The puborectalis muscle was considered dyssynergic when the ARA failed to open during straining *Proctogram:* Paradoxical puborectal muscle contraction was diagnosed as a persistent or exaggerated indentation of the puborectalis sling posteriorly at the anorectal junction without widening of the anorectal angle (ARA). The ARA was measured at the intersection of the axis of the anal canal and rectal ampulla	ARA failed to open during straining	N/A	ARA failed to open during straining
Martin 2017	*MRI* Thickening of the puborectalis muscle during prolonged evacuation of rectal contrast *EP* The anorectal angle (ARA) unchanged during defaecation in comparison with the angle at rest	Lack of opening of the anorectal angle during defaecation	Thickening of the puborectalis muscle during evacuation	N/A

Table 9. Anismus definition and cut-off value used in included studies *(Continued)*

Murad-Regadas 2008; Murad-Regadas 2011; Regadas 2011	The muscles failed to relax or contracted during defaecation. The anorectal angle sizes were compared between the resting and straining positions to determine the occurrence of normal relaxation or paradoxical contraction	The angle decreased by a minimum of 1 degree at Valsalva	N/A	The angle decreased by a minimum of 1 degree at Valsalva
Pilkington 2012	No emptying or evidence of puborectalis spasm	No rectal contrast evacuated or persistent puborectalis spasm	No rectal contrast evacuated or persistent puborectalis spasm	N/A
Poncelet 2017	Anismus referred to the failure to relax the anal sphincter or puborectalis muscle during defaecation, possibly due to dyssynergia of the anal sphincter or too narrow an anorectal angle	Failure to relax the puborectalis muscle	Failure to relax the puborectalis muscle	N/A
Van Gruting 2017	Anismus is a paradoxical pelvic floor contraction during attempts to evacuate. Anismus was present if a paradoxic contraction of the puborectalis muscle during straining was visualised or as a persistent impression of the puborectalis muscle on the posterior rectal wall	Paradoxical contraction of the puborectalis muscle during straining	Paradoxical contraction of the puborectalis muscle during straining	Paradoxical contraction of the puborectalis muscle during straining

N/A = not applicable; ARA = Anorectal angle

Table 10. Pelvic floor descent definition and cut-off value used in included studies

Study ID	Pelvic floor descent definition	EP - Pelvic floor descent cut-off value	MRI - Pelvic floor descent cut-off value	Ultrasound - Pelvic floor descent cut-off value
Barthet 2000	*DAE:* The descent of the puborectal muscle corresponded to the distance between its initial position and its position at the end of the straining effort *Defaecography:* Pelvic floor descent was estimated by comparing the level of the anorectal junction at rest and during straining	Anorectal angle > 2 cm below PCL at rest or > 3 cm below PCL on straining	N/A	Not described
Dellemare 1994	The anorectal junction was defined as the intersection point of the central axis of the anal canal and the line along the posterior wall of the distal rectum	The anorectal junction below tuber ischiadicum	The anorectal junction below baseline (cranial side symphysis and distal sacrum)	N/A
Fiaschetti 2013	Pelvic floor descent was measured by drawing a line perpendicular to the PCL from the posterior border of the H line. Abnormal descent occurred when the length of the M line was > 2 cm	M line was > 2 cm	M line was > 2 cm	N/A

Table 10. Pelvic floor descent definition and cut-off value used in included studies *(Continued)*

Foti 2013	The anorectal junction descends > 5 cm below the PCL during straining or defaecation	ARJ > 5 cm below PCL during straining	ARJ > 5 cm below PCL during straining	N/A
Martellucci 2011	Pelvic floor descent was defined as descent of the anorectal junction during straining by > 3.5 cm from its resting position at the inferior plane of the ischial tuberosities	ARJ movement ≥ 3.5 cm between rest and Valsalva	N/A	ARJ movement ≥ 3.5 cm between rest and Valsalva
Martin 2017	Distance between the anal margin and the sacropubic line with a 90 ° angle	Difference of ≥ 3.5 cm between the anorectal junction at straining and at rest	Anorectal junction > 30 mm below the PCL at Valsalva	N/A
Murad-Regadas 2011	*Ultrasound:* Pelvic floor descent was quantified by measuring the distance between the position of the proximal border of the PR at rest and the point to which it had been displaced by maximum straining (PR descent) *Defaecography:* Pelvid floor descent was considered a difference of > 3 cm in the position of the anal canal between relaxation and straining	> 3 cm difference of anal canal position between relaxation and straining	N/A	PR movement ≥ 2.5 cm between rest and Valsalva
Vanbeckevoort 1999	The diagnosis of descent of the rectum was based on measurement of the vertical distance between the pubococcygeal line and the anorectal junction	> 2.5 cm below the PCL	> 2.5 cm below the PCL	N/A
Van Gruting 2017	Pelvic floor descent is the abnormal descent of the pelvic floor and is assessed by measuring the position of the anorectal junction (ARJ) below the PCL at rest and during straining	ARJ > 30 mm below the PCL at Valsalva	ARJ > 30 mm below the PCL at Valsalva	N/A
Van Iersel 2017	The distance of the line perpendicular to the PCL to the cranial side of the anal canal (ARJ descent) during maximal straining or evacuation was measured in millimetres. An ARJ which descended by > 30 mm on straining signified excessive descent	ARJ > 30 mm below the PCL at Valsalva	ARJ > 30 mm below the PCL at Valsalva	N/A
Vitton 2011	*Fluoroscopy:* Pelvic floor descent was determined by measuring the level of the anorectal junction at rest and during straining and was defined as either descent of the anorectal angle to more than 2 cm below the pubococcygeal line at rest or descent to more than 3 cm below the pubococcygeal line on straining *MRI* Pelvic floor descent was defined as the descent of the anorectal junction to below the pubococcygeal line *DAE:*	ARJ > 3 cm below PCL on straining	ARJ below PCL	> 2 cm descent puborectalis muscle on straining

Table 10. Pelvic floor descent definition and cut-off value used in included studies *(Continued)*

	Pelvic floor descent was diagnosed when the distance between the initial and final positions of the descent of the puborectalis muscle was > 2 cm

N/A = not applicable; ARJ = Anorectal junction; PCL = Pubococcygeal line

Table 11. Main analysis: DTA characteristics for diagnosis of the five target conditions

Target condition	Imaging	Number of studies (participants)	Prevalence (%) (95% CrI)	Most used cut-off (%)	Diagnostic Odds Ratio (95% CrI)	Sensitivity (95% CrI)	Specificity (95% CrI)	PPV (95% CrI)	NPV (95% CrI)	LR+ (95% CrI)	LR- (95% CrI)
Rectocele	EP	34 (1737)	58.9 (51.3 - 67.8)	> 0 cm (44%) > 2 cm (35%)	143.7 (40.0 to 694.5)	0.98 (0.94 to 0.99)	0.78 (0.63 to 0.90)	0.86 (0.76 to 0.95)	0.96 (0.88 to 0.99)	4.37 (2.66 to 10.00)	0.03 (0.01 to 0.08)
	MRI	19 (659)		> 0 cm (37%) > 2 cm (32%)	160.4 (35.0 to 952.3)	0.94 (0.86 to 0.98)	0.90 (0.78 to 0.97)	0.93 (0.85 to 0.98)	0.92 (0.77 to 0.98)	9.63 (4.29 to 36.51)	0.06 (0.02 to 0.16)
	TPUS	11 (988)		> 1 cm (46%) > 2 cm (46%)	66.5 (17.8 to 322.0)	0.88 (0.75 to 0.97)	0.89 (0.81 to 0.96)	0.92 (0.85 to 0.97)	0.84 (0.66 to 0.95)	8.03 (4.37 to 21.48)	0.13 (0.04 to 0.29)
	EVUS	2 (454)		> 0 cm (50%) > 1 cm (50%)	7.6 (2.2 to 38.1)	0.69 (0.52 to 0.89)	0.76 (0.54 to 0.93)	0.81 (0.67 to 0.93)	0.63 (0.47 to 0.82)	2.91 (1.47 to 9.49)	0.41 (0.16 to 0.69)
	DAE	2 (99)		> 0 cm (50%) > 2 cm (50%)	23.8 (3.7 to 266.1)	0.75 (0.54 to 0.92)	0.88 (0.62 to 0.98)	0.90 (0.71 to 0.99)	0.70 (0.52 to 0.89)	6.38 (1.88 to 49.44)	0.29 (0.10 to 0.57)
	EDF	4 (169)		> 0 cm (100%)	231.3 (21.5 to 3691.7)	0.96 (0.87 to 0.99)	0.89 (0.60 to 0.99)	0.93 (0.76 to 0.99)	0.94 (0.80 to 0.99)	8.68 (2.33 to 75.52)	0.04 (0.01 to 0.16)
Enterocele	EP	31 (2233)	24.1 (19.6 - 28.7)	> 0 cm below PCL (32%)	295.9 (105.1 to 1520.3)	0.91 (0.83 to 0.97)	0.96 (0.93 to 0.99)	0.89 (0.80 to 0.97)	0.97 (0.94 to 0.99)	20.25.8 (13.7 to 81.6)	0.09 (0.03 to 0.17)

Table 11. Main analysis: DTA characteristics for diagnosis of the five target conditions *(Continued)*

				RV space (35%)							
	MRI	17 (1222)		> 0 cm below PCL (65%)	726.8 (129.6 to 5768.0)	0.85 (0.72 to 0.94)	0.99 (0.96 to 1.00)	0.97 (0.87 to 1.00)	0.95 (0.91 to 0.98)	110.0 (22.6 to 670.1)	0.16 (0.06 to 0.28)
	TPUS	10 (976)		RV septum (50%)	346.7 (60.7 to 3900.2)	0.84 (0.63 to 0.96)	0.98 (0.95 to 1.00)	0.94 (0.83 to 0.99)	0.95 (0.89 to 0.99)	52.2 (16.0 to 340.5)	0.17 (0.04 to 0.37)
	EVUS	3 (471)		RV septum (100%)	69.0 (9.6 to 460.2)	0.68 (0.51 to 0.91)	0.97 (0.80 to 0.99)	0.87 (0.53 to 0.97)	0.90 (0.85 to 0.97)	21.5 (3.6 to 89.1)	0.34 (0.09 to 0.51)
	DAE	2 (70)		RV space (50%)	94.4 (7.4 to 1294.5)	0.74 (0.52 to 0.94)	0.97 (0.75 to 1.00)	0.88 (0.48 to 0.98)	0.92 (0.85 to 0.98)	23.2 (2.9 to 187.2)	0.27 (0.06 to 0.51)
	EDF	3 (139)		Ischio-cochial line (66%)	102.0 (13.3 to 1544.2)	0.71 (0.51 to 0.96)	0.97 (0.87 to 1.00)	0.90 (0.62 to 0.98)	0.91 (0.85 to 0.99)	27.0 (5.3 to 174.4)	0.30 (0.04 to 0.51)
Intussusception	EP	27 (1613)	44.1 (34.7 - 52.6)	Any (70%) Full (30%	94.1 (33.1 to 433.3)	0.89 (0.79 to 0.96)	0.92 (0.86 to 0.97)	0.89 (0.80 to 0.97)	0.91 (0.82 to 0.97)	10.8 (6.2 to 31.0)	0.12 (0.04 to 0.23)
	MRI	12 (536)		Any (58%) Full (42%)	48.1 (10.8 to 387.5)	0.61 (0.51 to 0.78)	0.97 (0.88 to 1.00)	0.94 (0.77 to 0.99)	0.76 (0.66 to 0.87)	18.7 (5.1 to 131.6)	0.41 (0.23 to 0.52)
	TPUS	10 (664)		Any (80%) Full (20%)	87.3 (20.1 to 624.6)	0.75 (0.54 to 0.93)	0.96 (0.91 to 0.99)	0.94 (0.84 to 0.99)	0.83 (0.69 to 0.95)	20.8 (7.8 to 88.1)	0.26 (0.07 to 0.49)
	EVUS	2 (454)		Full (100%)	23.7	0.63	0.93	0.87	0.76	8.7 (2.2 to 50.7)	0.40

Table 11. Main analysis: DTA characteristics for diagnosis of the five target conditions *(Continued)*

					(4.1 to 197.2)	(0.51 to 0.88)	(0.72 to 0.99)	(0.62 to 0.98)	(0.65 to 0.91)		(0.14 to 0.57)
	DAE	2 (99)		Full (100%)	22.8 (2.9 to 264.6)	0.61 (0.50 to 0.89)	0.93 (0.65 to 0.99)	0.87 (0.57 to 0.98)	0.75 (0.63 to 0.92)	8.7 (1.8 to 64.8)	0.43 (0.12 to 0.64)
	EDF	4 (169)		Any (100%)	108.5 (12.7 to 1772.6)	0.89 (0.65 to 0.98)	0.92 (0.72 to 0.99)	0.90 (0.69 to 0.99)	0.92 (0.75 to 0.99)	11.4 (3.1 to 79.2)	0.12 (0.02 to 0.39)
Anismus	EP	15 (985)	24.8 (18.6 - 31.6)	ARA (53%) Paradox (33%)	132.5 (44.5 to 583.5)	0.80 (0.63 to 0.94)	0.97 (0.94 to 0.99)	0.89 (0.80 to 0.96)	0.94 (0.87 to 0.98)	25.17 (13.79 to 64.76)	0.20 (0.07 to 0.38)
	MRI	7 (287)		ARA (29%) Paradox (71%)	139.4 (25.7 to 1508.5)	0.86 (0.60 to 0.98)	0.96 (0.89 to 0.99)	0.87 (0.69 to 0.96)	0.95 (0.86 to 0.99)	19.59 (7.58 to 61.01)	0.15 (0.02 to 0.42)
	TPUS	5 (651)		ARA (20%) Paradox (40%)	123.5 (24.0 to 741.9)	0.92 (0.72 to 0.98)	0.91 (0.83 to 0.97)	0.77 (0.61 to 0.91)	0.97 (0.9 to 0.99)	10.38 (5.24 to 26.91)	0.09 (0.02 to 0.30)
	EVUS	2 (454)		ARA (50%) Paradox (50%)	52.6 (5.8 to 384.4)	0.84 (0.59 to 0.96)	0.90 (0.63 to 0.98)	0.74 (0.40 to 0.92)	0.95 (0.86 to 0.99)	8.72 (2.12 to 34.70)	0.18 (0.04 to 0.49)
	DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	4 (169)		ARA (100%)	95.9 (13.8 to 1105.1)	0.87 (0.72 to 0.96)	0.93 (0.74 to 0.99)	0.80 (0.49 to 0.97)	0.96 (0.90 to 0.99)	12.1 (3.24 to 94.67)	0.14 (0.04 to 0.32)
Pelvic floor descent	EP	10 (476)	66.9 (55.0 - 78.1)	> 3 cm ARJ below PCL (50%)	207.9 (34.2 to 2119.6)	0.98 (0.93 to 1.00)	0.83 (0.59 to 0.96)	0.92 (0.79 to 0.98)	0.94 (0.79 to 0.99)	5.83 (2.34 to 25.77)	0.03 (0.01 to 0.1)

Table 11. Main analysis: DTA characteristics for diagnosis of the five target conditions *(Continued)*

MRI	7 (350)	> 3 cm ARJ below PCL (43%)	59.2 (11.3 to 684.6)	0.94 (0.81 to 0.98)	0.79 (0.54 to 0.97)	0.90 (0.76 to 0.99)	0.86 (0.61 to 0.97)	4.46 (2.00 to 28.23)	0.08 (0.02 to 0.24)
TPUS	1 (54)	> 3.5 cm diff ARJ rest-Valsalva (100%)	140.3 (5.8 to 3770.6)	0.88 (0.55 to 0.99)	0.95 (0.62 to 1.00)	0.97 (0.80 to 1.00)	0.78 (0.46 to 0.97)	17.19 (2.12 to 179.03)	0.14 (0.01 to 0.50)
EVUS	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	2 (99)	> 2 cm PR movement (50%)	41.0 (4.3 to 492.5)	0.93 (0.64 to 0.99)	0.74 (0.54 to 0.93)	0.88 (0.75 to 0.97)	0.84 (0.47 to 0.98)	3.5 (1.85 to 13.54)	0.1 (0.01 to 0.5)
EDF	1 (29)	> 2.5 cm PR movement (50%)	2.7 (4.7 to 1858.7)	0.85 (0.55 to 0.98)	0.93 (0.60 to 0.99)	0.96 (0.79 to 1.00)	0.74 (0.43 to 0.96)	10.96 (1.99 to 110.4)	0.17 (0.02 to 0.53)

For abbreviations and overview of all used cut-off values see Table 6; Table 7; Table 8; Table 9; Table 10.
Numbers provided are median (95% CrI) unless otherwise stated; CrI = credibility Interval; PPV = Positive predictive value; NPV = Negative predictive value; LR+ = Likelihood ratio positive; LR- = Likelihood ratio negative

Table 12. Main analysis: Probability that index test is equal or better than EP

Target condition	Index test	Pooled sensitivity			Pooled specificity		
		Estimate (%) (95% CrI)	Difference vs EP (%)	Probability	Estimate (%) (95% CrI)	Difference vs EP (%)	Probability
Rectocele	**EP**	97.5 (93.7 to 99.3)	N/A	N/A	77.8 (63.5 to 90.2)	N/A	N/A
	MRI	94.3 (85.9 to 98.4)	−3.1 (−11.8 to 2.2)	0.127	90.3 (78.5 to 97.4)	12.2 (−4.5 to 28.6)	0.924
	TPUS	88.4 (74.8 to 96.6)	−8.9 (−22.7 to −0.1)	0.024	89.1 (80.8 to 95.9)	11.2 (−3.7 to 26.9)	0.929
	EVUS	69.0 (51.5 to 88.8)	−28.3 (−45.9 to −8.4)	0.001	76.5 (53.5 to 92.9)	−1.4 (−27.4 to 21.1)	0.454
	DAE	74.6 (53.8 to 91.6)	−22.6 (−43.6 to −5.4)	0.002	88.5 (61.6 to 98.5)	10.1 (−17.3 to 28.3)	0.786

Table 12. Main analysis: Probability that index test is equal or better than EP *(Continued)*

	EDF	96.4 (86.8 to 99.4)	−1.1 (−10.8 to 3.6)	0.329	89.0 (59.7 to 98.7)	10.3 (−19.8 to 29.4)	0.804
Enterocele	**EP**	91.2 (83.2 to 97.1)	N/A	N/A	96.5 (93.4 to 98.9)	N/A	N/A
	MRI	84.5 (71.8 to 94.0)	−6.6 (−20.5 to 5.9)	0.152	99.2 (96.3 to 99.9)	2.6 (−1.2 to 5.9)	0.931
	TPUS	83.6 (63.1 to 96.0)	−7.5 (−28.5 to 7.0)	0.180	98.4 (95.1 to 99.8)	1.8 (−2.2 to 5.2)	0.842
	EVUS	67.7 (51.2 to 91.4)	−23.2 (−41.8 to 1.4)	0.035	96.9 (80.2 to 99.2)	0.2 (−16.2 to 4.4)	0.527
	DAE	74.5 (52.4 to 94.3)	−16.6 (−39.5 to 4.4)	0.079	96.8 (75.2 to 99.6)	0.1 (−21.1 to 4.7)	0.518
	EDF	70.9 (51.2 to 95.9)	−20.2 (−41.2 to 6.1)	0.092	97.4 (86.9 to 99.6)	0.8 (−9.9 to 4.7)	0.618
Intussusception	**EP**	88.8 (78.8 to 96.3)	N/A	N/A	91.8 (85.9 to 97.2)	N/A	N/A
	MRI	60.6 (50.8 to 78.1)	−27.8 (−41.1 to −8.3)	0.003	96.7 (88.1 to 99.5)	4.5 (−4.9 to 11.5)	0.865
	TPUS	75.0 (53.6 to 92.8)	−13.6 (−36.2 to 6.5)	0.104	96.4 (90.9 to 99.1)	4.4 (−2.9 to 11.1)	0.891
	EVUS	63.2 (51.1 to 87.5)	−25.1 (−40.4 to −0.7)	0.022	92.6 (71.5 to 98.7)	0.6 (−20.4 to 9.3)	0.546
	DAE	61.4 (50.5 to 89.2)	−26.7 (−42.0 to 1.7)	0.035	92.7 (64.6 to 99.0)	0.6 (−27.9 to 10.1)	0.539
	EDF	89.3 (65.1 to 98.5)	0.4 (−24.4 to 14.6)	0.517	92.4 (71.9 to 98.9)	0.3 (−19.9 to 9.7)	0.518
Anismus	**EP**	80.4 (63.1 to 93.7)	N/A	N/A	96.8 (94.4 to 98.8)	N/A	N/A
	MRI	85.9 (60.4 to 98.2)	4.7 (−20.3 to 25.7)	0.656	95.8 (89.4 to 98.6)	−1.1 (−7.6 to 2.7)	0.291
	TPUS	91.9 (72.1 to 98.3)	10.7 (−10.5 to 29.4)	0.864	91.3 (83.1 to 96.7)	−5.5 (−13.8 to 0.2)	0.030
	EVUS	84.5 (59.1 to 96.2)	3.7 (−23.9 to 25.3)	0.617	90.5 (63.0 to 97.6)	−6.4 (−34.1 to 1.1)	0.061
	DAE	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	87.3 (71.6 to 96.2)	6.5 (−13.2 to 26.2)	0.751	92.9 (73.8 to 99.1)	−3.9 (−23.1 to 2.9)	0.195
PFD	**EP**	97.5 (92.6 to 99.5)	N/A	N/A	83.3 (58.7 to 96.2)	N/A	N/A
	MRI	93.8 (81.4 to 98.4)	−3.6 (−15.8 to 2.8)	0.127	79.2 (53.7 to 96.7)	−3.3 (−33.8 to 26.2)	0.416

Table 12. Main analysis: Probability that index test is equal or better than EP *(Continued)*

TPUS	87.5 (55.4 to 98.7)	−9.8 (2.4 to −42.0)	0.094	95.1 (61.8 to 99.5)	10 (36 to −22.5)	0.809
EVUS	N/A	N/A	N/A	N/A	N/A	N/A
DAE	92.9 (64.4 to 99.1)	−4.4 (−32.9 to 3)	0.162	74.2 (53.6 to 93.4)	−8.1 (−33.4 to 20 to)	0.282
EDF	84.5 (55.0 to 98.3)	−12.7 (−42.5 to 1.6)	0.058	92.6 (60.0 to 99.3)	7.9 (−25.3 to 34.5)	0.748

Probability of < 0.400 means estimated test accuracy of index test is not equal or higher than EP
Probability of > 0.400 means estimated test accuracy of index test is equal or higher than EP

Table 13. Heterogeneity analysis: TPUS with and without rectal contrast

Target condition	Rectal Contrast Yes			Rectal Contrast No			Probability *Rectal contrast better than no rectal contrast*	
	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	Sensitivity	Specificity
Rectocele	3 (240)	0.92 (0.69 to 0.99)	0.87 (0.69 to 0.97)	7 (694)	0.81 (0.58 to 0.95)	0.88 (0.73 to 0.96)	0.821	0.451
Enterocele	4 (302)	0.90 (0.71 to 0.99)	0.95 (0.86 to 0.99)	5 (620)	0.67 (0.51 to 0.90)	0.99 (0.92 to 1.00)	0.936	0.128
Intussusception	3 (240)	0.90 (0.69 to 0.98)	0.90 (0.68 to 0.99)	6 (370)	0.61 (0.51 to 0.86)	0.96 (0.89 to 0.99)	0.967	0.189
Anismus	1 (102)	N/A	N/A	3 (495)	N/A	N/A	N/A	N/A
Pelvic floor descent	0 (0)	N/A	N/A	0 (0)	N/A	N/A	N/A	N/A

Probability of < 0.400 means estimated test accuracy of rectal contrast is not equal or higher than no rectal contrast
Probability of > 0.400 means estimated test accuracy of rectal contrast is equal or higher than no rectal contrast
N/A = not analysable

Table 14. Heterogeneity analysis: MRI with and without evacuation phase

Target condition	Evacuation Phase Yes			Evacuation Phase No			Probability *Evacuation phase better than no evacuation phase*	
	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	Sensitivity	Specificity
Rectocele	13 (572)	0.94 (0.87 to 0.98)	0.84 (0.67 to 0.95)	7 (104)	0.65 (0.51 to 0.89)	0.95 (0.84 to 1.00)	0.991	0.072
Enterocele	13 (1159)	0.87 (0.74 to 0.95)	0.99 (0.94 to 1.00)	5 (80)	0.62 (0.51 to 0.88)	0.97 (0.88 to 1.00)	0.954	0.759
Intussusception	11 (527)	0.63 (0.51 to 0.81)	0.96 (0.86 to 1.00)	2 (26)	0.59 (0.50 to 0.86)	0.93 (0.62 to 0.99)	0.594	0.669
Anismus	6 (277)	0.84 (0.58 to 0.99)	0.96 (0.89 to 0.99)	2 (27)	0.89 (0.57 to 0.99)	0.93 (0.66 to 0.99)	0.391	0.677
Pelvic floor descent	6 (315)	N/A	N/A	2 (52)	N/A	N/A	N/A	N/A

Probability of < 0.400 means estimated test accuracy of evacuation phase is not equal or higher than no evacuation phase
Probability of > 0.400 means estimated test accuracy of evacuation phase is equal or higher than no evacuation phase
N/A = not analysable

Table 15. Heterogeneity analysis: Cut-off values

Target condition	Imaging technique	Cut-off A				Cut-off B				Probability *Cut-off B better than cut-off A*	
		Cut-off value	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	Cut-off value	# Studies (# participants)	Sensitivity (95% CrI)	Specificity (95% CrI)	Sensitivity	Specificity
Rectocele	**EP**	> 0 cm depth	19 (737)	0.99 (0.98 to 1.00)	0.55 (0.50 to 0.67)	≥ 2 cm depth	27 (1410)	0.97 (0.91 to 0.99)	0.89 (0.78 to 0.96)	0.034	1.000

Table 15. Heterogeneity analysis: Cut-off values *(Continued)*

	MRI		10 (313)	0.98 (0.92 to 1.00)	0.66 (0.51 to 0.86)		14 (575)	0.93 (0.83 to 0.98)	0.94 (0.85 to 0.99)	0.092	0.995
	TPUS		7 (418)	0.83 (0.59 to 0.95)	0.69 (0.51 to 0.90)		7 (742)	0.91 (0.74 to 0.99)	0.91 (0.79 to 0.97)	0.806	0.956
Entero-cele	**EP**	> 0 cm below PCL	15 (502)	0.90 (0.74 to 0.98)	0.96 (0.91 to 0.99)	Recto-vaginal Space/ septum	15 (1624)	0.90 (0.79 to 0.97)	0.95 (0.91 to 0.98)	0.505	0.300
	MRI		12 (473)	0.84 (0.69 to 0.94)	0.99 (0.94 to 1.00)		3 (702)	0.93 (0.71 to 0.99)	0.98 (0.76 to 1.00)	0.801	0.392
	TPUS		0 (0)	N/A	N/A		5 (344)	N/A	N/A	N/A	N/A
Intus-suscep-tion	**EP**	Any	21 (1052)	0.93 (0.82 to 0.99)	0.90 (0.81 to 0.97)	Full thick-ness cir-cumfer-ential	12 (692)	0.83 (0.81 to 0.98)	0.87 (0.65 to 0.98)	0.128	0.374
	MRI		8 (402)	0.58 (0.50 to 0.76)	0.92 (0.77 to 0.98)		5 (256)	0.70 (0.52 to 0.91)	0.92 (0.75 to 0.99)	0.803	0.526
	TPUS		9 (634)	0.88 (0.60 to 0.98)	0.95 (0.88 to 0.99)		2 (161)	0.61 (0.50 to 0.89)	0.95 (0.63 to 1.00)	0.071	0.503
Anismus	**EP**	Paradox-ical con-traction	8 (554)	0.55 (0.50 to 0.75)	0.96 (0.89 to 0.99)	ARA more acute	8 (288)	0.85 (0.70 to 0.97)	0.98 (0.94 to 1.00)	0.988	0.875
	MRI		4 (222)	0.70 (0.52 to 0.95)	0.95 (0.83 to 0.99)		3 (65)	0.91 (0.62 to 0.99)	0.96 (0.80 to 0.99)	0.858	0.569
	TPUS		2 (454)	N/A	N/A		1 (54)	N/A	N/A	N/A	N/A
Pelvic floor de-scent	**EP**	> 0 cm below PCL	6 (309)	N/A	N/A	Differ-ence ARJ rest to valsalva	3 (121)	N/A	N/A	N/A	N/A
	MRI		6 (301)	N/A	N/A		0 (0)	N/A	N/A	N/A	N/A

PCL = pubococcygeal line
ARA = Anorectal angle
ARJ = Anorectal juction
Probability of < 0.400 means estimated test accuracy of cutto off B is not equal or higher than cut-off A
Probability of > 0.400 means estimated test accuracy of cut-off B is equal or higher than cut-off A
N/A = not analysable

Table 16. Sensitivity analysis 1: DTA characteristics excluding studies with high risk of bias

Target condition	Imaging	Number of studies (partici-pants)	Preva-lence (%) (95% CI)	Diagnostic odds ratio (95% CrI)	Sensitivity (95% CrI)	Specificity (95% CrI)	PPV (95% CrI)	NPV (95% CrI)	LR+ (95% CrI)	LR− (95% CrI)
Rectocele	**EP**	24 (1482)	59.7 (50.8 to 68.9)	126.8 (37.8 to 580.7)	0.97 (0.92 to 0.99)	0.80 (0.68 to 0.89)	0.88 (0.78 to 0.95)	0.95 (0.85 to 0.99)	4.72 (3.01 to 9.08)	0.04 (0.01 to 0.10)
	MRI	12 (504)		95.4 (16.5 to 638.7)	0.93 (0.78 to 0.98)	0.88 (0.72 to 0.97)	0.92 (0.82 to 0.98)	0.89 (0.67 to 0.97)	7.38 (3.22 to 29.98)	0.08 (0.02 to 0.26)
	TPUS	10 (947)		89.7 (23.7 to 494.7)	0.91 (0.80 to 0.97)	0.90 (0.80 to 0.97)	0.93 (0.86 to 0.98)	0.87 (0.7 to 0.96)	8.93 (4.46 to 26.36)	0.11 (0.03 to 0.23)
	EVUS	2 (454)		8.1 (2.2 to 36.9)	0.70 (0.52 to 0.88)	0.77 (0.54 to 0.93)	0.82 (0.66 to 0.94)	0.63 (0.46 to 0.82)	2.99 (1.48 to 9.52)	0.40 (0.16 to 0.69)
	DAE	2 (99)		25.6 (3.9 to 261.8)	0.74 (0.54 to 0.91)	0.90 (0.63 to 0.99)	0.91 (0.73 to 0.99)	0.69 (0.51 to 0.88)	7.05 (1.96 to 49.47)	0.30 (0.10 to 0.56)
	EDF	2 (110)		56.6 (5.4 to 1223.3)	0.93 (0.68 to 0.99)	0.81 (0.52 to 0.98)	0.88 (0.71 to 0.99)	0.88 (0.59 to 0.98)	4.69 (1.85 to 50.56)	0.09 (0.02 to 0.43)
Entero-cele	**EP**	22 (1996)	21.4 (17.4 to 25.9)	349.1 (115.9 to 1632)	0.94 (0.86 to 0.98)	0.96 (0.93 to 0.99)	0.86 (0.75 to 0.95)	0.98 (0.96 to 1)	22.14 (12.88 to 62.12)	0.07 (0.02 to 0.15)
	MRI	11 (1056)		398.9 (88.6 to 2341)	0.73	0.99 (0.97 to 1.00)	0.97	0.93	105.22	0.27

Table 16. Sensitivity analysis 1: DTA characteristics excluding studies with high risk of bias *(Continued)*

					(0.61 to 0.86)		(0.87 to 0.99)	(0.89 to 0.97)	(26.37 to 538.34)	(0.15 to 0.40)
	TPUS	9 (935)		277.2 (49.7 to 2338.1)	0.86 (0.66 to 0.97)	0.98 (0.93 to 1.00)	0.91 (0.76 to 0.98)	0.96 (0.91 to 0.99)	36.6 (12.27 to 165.17)	0.15 (0.03 to 0.35)
	EVUS	3 (471)		65.2 (9.2 to 462.3)	0.68 (0.51 to 0.92)	0.97 (0.80 to 0.99)	0.85 (0.49 to 0.96)	0.92 (0.87 to 0.98)	20.32 (3.65 to 84.86)	0.34 (0.09 to 0.51)
	DAE	1 (56)		41.3 (2.9 to 581.2)	0.66 (0.51 to 0.90)	0.95 (0.62 to 1.00)	0.79 (0.32 to 0.97)	0.91 (0.85 to 0.97)	13.64 (1.75 to 134.88)	0.37 (0.11 to 0.62)
	EDF	2 (110)		87.0 (7.4 to 1365.9)	0.69 (0.51 to 0.95)	0.97 (0.79 to 1.00)	0.87 (0.46 to 0.98)	0.92 (0.86 to 0.99)	24.89 (3.23 to 169.12)	0.33 (0.05 to 0.52)
Intussus-ception	**EP**	22 (1462)	45.5 (36.2 to 54.2)	101.2 (33.4 to 494.8)	0.91 (0.81 to 0.97)	0.91 (0.85 to 0.96)	0.89 (0.8 to 0.96)	0.92 (0.83 to 0.98)	9.67 (5.91 to 25.26)	0.10 (0.03 to 0.21)
	MRI	10 (485)		58.5 (10.4 to 481.5)	0.62 (0.51 to 0.81)	0.97 (0.88 to 1.00)	0.95 (0.78 to 0.99)	0.75 (0.65 to 0.87)	21.87 (4.91 to 157.37)	0.40 (0.20 to 0.52)
	TPUS	9 (632)		85.9 (19.7 to 630.4)	0.73 (0.52 to 0.92)	0.97 (0.91 to 0.99)	0.95 (0.85 to 0.99)	0.81 (0.67 to 0.94)	22.4 (7.99 to 99.89)	0.28 (0.08 to 0.50)
	EVUS	2 (454)		23.1 (4.6 to 190.5)	0.64 (051 to 0.86)	0.93 (0.73 to 0.99)	0.88 (0.65 to 0.98)	0.75 (0.64 to 0.9)	8.63 (2.34 to 49.9)	0.40 (0.15 to 0.56)
	DAE	2 (99)		22.3 (2.9 to 254.6)	0.62	0.93 (0.64 to 0.99)	0.87	0.75	8.48 (1.75 to 60.53)	0.42

Table 16. Sensitivity analysis 1: DTA characteristics excluding studies with high risk of bias *(Continued)*

					(0.51 to 0.89)		(0.58 to 0.98)	(0.62 to 0.91)		(0.12 to 0.64)
	EDF	2 (110)		93.5 (7.6 to 1533.7)	0.91 (0.66 to 0.99)	0.90 (0.62 to 0. 99)	0.88 (0.64 to 0.98)	0.92 (0.73 to 0.99)	8.56 (2.24 to 59.03)	0.10 (0.01 to 0.42)
Anismus	**EP**	9 (825)	24.8 (17.4 to 33.9)	48.9 (20.2 to 173.9)	0.63 (0.52 to 0.83)	0.97 (0.93 to 0.97)	0.86 (0.72 to 0.94)	0.89 (0.81 to 0.96)	18.03 (9.09 to 45.1)	0.38 (0.18 to 0.50)
	MRI	4 (227)		70.6 (11.9 to 801)	0.76 (0.53 to 0.96)	0.96 (0.84 to 0.99)	0.85 (0.58 to 0.96)	0.92 (0.83 to 0.99)	16.98 (4.63 to 63.48)	0.26 (0.04 to 0.50)
	TPUS	4 (610)		98.5 (16.5 to 691)	0.90 (0.64 to 0.98)	0.91 (0.81 to 0.97)	0.77 (0.56 to 0.91)	0.97 (0.87 to 0.99)	10.07 (4.47 to 27.65)	0.11 (0.02 to 0.40)
	EVUS	2 (454)		60.0 (6.0 to 449.5)	0.86 (0.60 to 0.97)	0.91 (0.63 to 0.98)	0.75 (0.4 to 0.93)	0.95 (0.85 to 0.99)	9.21 (2.13 to 35.21)	0.16 (0.04 to 0.48)
	DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	2 (110)		52.2 (5.2 to 767.3)	0.84 (0.59 to 0.96)	0.90 (0.60 to 0.99)	0.74 (0.38 to 0.96)	0.94 (0.84 to 0.99)	8.51 (2.02 to 78.44)	0.18 (0.04 to 0.49)
Pelvic floor descent	**EP**	7 (374)	71.0 (58.2 to 82.4)	114.5 (20.5 to 895.3)	0.96 (0.91 to 0.99)	0.81 (0.54 to 0.96)	0.92 (0.8 to 0.99)	0.90 (0.72 to 0.97)	5 (2.09 to 22.31)	0.05 (0.01 to 0.13)
	MRI	5 (277)		159.0 (15.5 to 2325.6)	0.94 (0.73 to 0.99)	0.91 (0.63 to 0.99)	0.96 (0.83 to 1)	0.86 (0.5 to 0.97)	9.92 (2.53 to 91.6)	0.07 (0.02 to 0.31)
	TPUS	1 (54)		115.3	0.87	0.95	0.97	0.74	15.49	0.15

Table 16. Sensitivity analysis 1: DTA characteristics excluding studies with high risk of bias *(Continued)*

		(5.4 to 3162.5)	(0.55 to 0.99)	(0.62 to 0.99)	(0.83 to 1)	(0.4 to 0.97)	(2.08 to 138.5)	(0.02 to 0.50)
EVUS	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	2 (99)	38.3 (3.9 to 508.4)	0.92 (0.64 to 0.99)	0.75 (0.54 to 0.95)	0.90 (0.77 to 0.98)	0.79 (0.4 to 0.97)	3.6 (1.81 to 16.95)	0.11 (0.01 to 0.51)
EDF	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 17. Sensitivity analysis 1: Probability that index test is equal or better than EP

Target condition	Index test	Pooled sensitivity			Pooled specificity		
		Estimate (%) (95% CrI)	Difference vs EP (%)	Probability	Estimate (%) (95% CrI)	Difference vs EP (%)	Probability
Rectocele	**EP**	96.9 (91.8 to 99.2)	N/A	N/A	79.5 (67.9 to 89.4)	N/A	N/A
	MRI	92.8 (78.1 to 98.0)	−4 (−18.6 to 3.1)	0.137	87.6 (72.4 to 96.9)	7.9 (−9.6 to 22.9)	0.826
	TPUS	90.5 (79.7 to 97.4)	−6.2 (−17.2 to 2)	0.075	89.9 (80.3 to 96.5)	10.3 (−3.2 to 23.3)	0.936
	EVUS	69.6 (51.8 to 88.3)	−27 (−45.2 to −7.7)	0.001	77.0 (53.7 to 93.1)	−2.6 (−27.3 to 17.4)	0.413
	DAE	73.9 (53.6 to 91.3)	−22.7 (−43.1 to −4.8)	0.003	89.6 (63.3 to 98.5)	9.6 (−17 to 24.5)	0.810
	EDF	92.7 (67.7 to 98.7)	−4 (−29 to 3.5)	0.175	80.7 (52.4 to 98.2)	1.2 (−29.1 to 23.2)	0.528
Enterocele	**EP**	93.5 (86 to 98.2)	N/A	N/A	95.8 (92.8 to 98.5)	N/A	N/A
	MRI	73.3 (60.5 to 85.5)	−20 (−33.5 to −5.9)	0.003	99.3 (97.2 to 99.9)	3.4 (0.2 to 6.6)	0.982
	TPUS	85.8 (65.6 to 97.0)	−7.5 (−28.1 to 5.6)	0.158	97.7 (93.4 to 99.5)	1.8 (−3.1 to 5.5)	0.798
	EVUS	68.0 (51.2 to 91.6)	−24.9 (−43.3 to −1)	0.019	96.7 (80.2 to 99.2)	0.6 (−15.5 to 5.1)	0.573
	DAE	65.6 (50.8 to 90.1)	−27.5 (−44.1 to −2.2)	0.014	95.2 (62 to 99.5)	−0.7 (−33.7 to 5.1)	0.440

Table 17. Sensitivity analysis 1: Probability that index test is equal or better than EP *(Continued)*

	EDF	68.8 (51.1 to 95.3)	−24.2 (−43.4 to 2.6)	0.051	97.2 (79.3 to 99.6)	1.1 (−16.6 to 5.5)	0.645
Intussusception	**EP**	90.8 (81.2 to 97.2)	N/A	N/A	90.6 (84.9 to 96.4)	N/A	N/A
	MRI	61.5 (50.7 to 80.7)	−28.8 (−42.7 to −8.1)	0.003	97.1 (87.7 to 99.6)	6.1 (−4.3 to 12.7)	0.904
	TPUS	73.0 (52.4 to 91.9)	−17.5 (−39.2 to 3.4)	0.057	96.8 (91.4 to 99.3)	5.9 (−1.6 to 12.2)	0.947
	EVUS	63.6 (51.3 to 86.2)	−26.8 (−41.5 to -2.7)	0.014	92.6 (72.7 to 98.7)	1.7 (−18.3 to 10.6)	0.628
	DAE	61.8 (50.5 to 89.2)	−28.2 (−43.4 to 0.1)	0.025	92.6 (63.8 to 98.9)	1.5 (−27.2 to 11)	0.581
	EDF	91.1 (65.7 to 98.8)	0.1 (−25.6 to 13)	0.503	89.6 (62.0 to 98.5)	−1.4 (−28.8 to 9.7)	0.429
Anismus	**EP**	63.0 (51.6 to 83.2)	N/A	N/A	96.5 (93.2 to 98.6)	N/A	N/A
	MRI	75.5 (52.7 to 96.3)	11.2 (−15.1 to 35.2)	0.787	95.6 (84.4 to 98.8)	−0.9 (−11.8 to 3.8)	0.365
	TPUS	90.3 (63.8 to 98.2)	25.3 (−3.5 to 42)	0.960	91.3 (80.6 to 96.8)	−5.1 (−15.9 to 1.2)	0.057
	EVUS	85.5 (59.5 to 96.6)	21.3 (−8.9 to 39.1)	0.924	90.9 (62.5 to 97.6)	−5.6 (−33.7 to 1.9)	0.092
	DAE	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	84.0 (58.9 to 96.3)	19.3 (−8.6 to 38.1)	0.923	90.3 (60.4 to 98.9)	−6 (−35.9 to 3.2)	0.161
PFD	**EP**	96.3 (90.9 to 98.9)	N/A	N/A	80.8 (54.2 to 95.7)	N/A	N/A
	MRI	94.2 (72.7 to 98.7)	−2.1 (−22.9 to 4.5)	0.273	90.8 (63.4 to 99.0)	9.2 (−20.6 to 37.7)	0.768
	TPUS	86.8 (55.1 to 98.6)	−9.3 (−40.6 to 3.7)	0.135	94.6 (61.5 to 99.4)	11.7 (−20.1 to 40.3)	0.817
	EVUS	N/A	N/A	N/A	N/A	N/A	N/A
	DAE	92.0 (64.1 to 99.1)	−4.1 (−31.8 to 4.5)	0.229	75.1 (53.5 to 94.7)	−4.6 (−33.1 to 25.8)	0.388
	EDF	N/A	N/A	N/A	N/A	N/A	N/A

Probability of < 0.400 means estimated test accuracy of index test is not equal or higher than EP
Probability of > 0.400 means estimated test accuracy of index test is equal or higher than EP

Table 18. Sensitivity analysis 2: DTA characteristics in women with ODS only

Target condition	Imaging	Number of studies (partici-pants)	Preva-lence (%)(95% CI)	Diagnostic Odds Ratio (95% CrI)	Sensitivity (95% CrI)	Specificity (95% CrI)	PPV (95% CrI)	NPV (95% CrI)	LR+ (95% CrI)	LR- (95% CrI)
Rectocele	**EP**	22 (901)	60.7 (46.9 to 73.1)	146.4 (25.8 to 1075.6)	0.97 (0.91 to 0.99)	0.82 (0.62 to 0.94)	0.89 (0.72 to 0.97)	0.94 (0.83 to 0.99)	5.32 (2.48 to 16.53)	0.04 (0.01 to 0.12)
	MRI	11 (432)		47.3 (6.2 to 398.9)	0.89 (0.66 to 0.98)	0.84 (0.66 to 0.96)	0.90 (0.77 to 0.97)	0.83 (0.52 to 0.97)	5.62 (2.33 to 22.31)	0.13 (0.03 to 0.43)
	TPUS	7 (379)		22.5 (4.9 to 156.8)	0.79 (0.56 to 0.95)	0.85 (0.73 to 0.94)	0.89 (0.78 to 0.96)	0.73 (0.46 to 0.94)	5.19 (2.52 to 13.8)	0.24 (0.06 to 0.54)
	EVUS	1 (131)		12.5 (2 to 103.7)	0.60 (0.50 to 0.88)	0.89 (0.57 to 0.97)	0.89 (0.67 to 0.98)	0.60 (0.41 to 0.83)	5.31 (1.42 to 25.01)	0.46 (0.14 to 0.73)
	DAE	2 (99)		24.2 (3.7 to 289.2)	0.73 (0.53 to 0.91)	0.90 (0.63 to 0.99)	0.92 (0.71 to 0.99)	0.68 (0.46 to 0.89)	6.92 (1.88 to 53.43)	0.31 (0.10 to 0.57)
	EDF	4 (169)		210.4 (19.3 to 3225.7)	0.96 (0.85 to 0.99)	0.89 (0.60 to 0.99)	0.93 (0.76 to 0.99)	0.93 (0.75 to 0.99)	8.69 (2.37 to 69.64)	0.05 (0.01 to 0.18)
Entero-cele	**EP**	18 (787)	25.1 (19.6 to 30.8)	356.7 (109.4 to 1971.8)	0.90 (0.81 to 0.96)	0.98 (0.94 to 0.99)	0.92 (0.83 to 0.98)	0.97 (0.93 to 0.99)	35.61 (15.95 to 142.4)	0.11 (0.04 to 0.19)
	MRI	7 (342)		119.8 (32 to 616)	0.78 (0.60 to 0.90)	0.97 (0.92 to 0.99)	0.90 (0.75 to 0.97)	0.93 (0.87 to 0.97)	26.86 (9.78 to 106.2)	0.23 (0.1 to 0.41)
	TPUS	7 (410)		197.3	0.81	0.98	0.92	0.94	35.22	0.20

Table 18. Sensitivity analysis 2: DTA characteristics in women with ODS only *(Continued)*

				(32 to 1920.8)	(0.59 to 0.95)	(0.92 to 1.00)	0.74 to 0.99)	(0.87 to 0.99)	(9.24 to 204.39)	(0.05 to 0.42)
	EVUS	2 (148)		54.7 (5.7 to 506.5)	0.77 (0.54 to 0.95)	0.94 (0.68 to 0.99)	0.81 (0.43 to 0.96)	0.92 (0.84 to 0.98)	12.55 (2.36 to 61.18)	0.25 (0.05 to 0.52)
	DAE	2 (70)		94.1 (7.3 to 1399.1)	0.74 (0.53 to 0.94)	0.97 (0.76 to 1.00)	0.89 (0.49 to 0.99)	0.92 (0.84 to 0.98)	23.09 (2.9 to 200.08)	0.27 (0.06 to 0.51)
	EDF	3 (139)		96.5 (12.9 to 1286.5)	0.68 (0.51 to 0.94)	0.98 (0.88 to 1.00)	0.90 (0.63 to 0.98)	0.90 (0.84 to 0.98)	28.42 (5.31 to 171.28)	0.33 (0.06 to 0.51)
Intussus-ception	**EP**	20 (876)	41.3 (31.2 to 50.8)	86.7 (27.6 to 490.6)	0.88 (0.77 to 0.96)	0.92 (0.85 to 0.98)	0.88 (0.76 to 0.97)	0.91 (0.82 to 0.98)	10.84 (5.78 to 38.71)	0.14 (0.04 to 0.25)
	MRI	9 (408)		42.9 (9.4 to 328)	0.58 (0.50 to 0.78)	0.97 (0.88 to 1.00)	0.93 (0.74 to 0.99)	0.77 (0.67 to 0.88)	17.45 (4.67 to 118.76)	0.44 (0.23 to 0.53)
	TPUS	7 (378)		61.5 (13.8 to 477.7)	0.75 (0.53 to 0.94)	0.95 (0.88 to 0.99)	0.91 (0.78 to 0.98)	0.85 (0.71 to 0.96)	14.75 (5.85 to 59.7)	0.26 (0.07 to 0.50)
	EVUS	1 (131)		22.2 (2.7 to 296)	0.68 (0.51 to 0.95)	0.90 (0.59 to 0.99)	0.83 (0.52 to 0.97)	0.80 (0.65 to 0.96)	7.05 (1.64 to 46.02)	0.37 (0.06 to 0.64)
	DAE	2 (99)		22.3 (2.8 to 289.9)	0.62 (0.50 to 0.90)	0.93 (0.63 to 0.99)	0.86 (0.53 to 0.98)	0.78 (0.65 to 0.93)	8.55 (1.71 to 65.82)	0.43 (0.11 to 0.64)
	EDF	4 (169)		105.6	0.89 (0.64 to 0.99)	0.92 (0.72 to 0.99)	0.89	0.92	11.23 (2.97 to 91.05)	0.12 (0.02 to 0.4)

Table 18. Sensitivity analysis 2: DTA characteristics in women with ODS only *(Continued)*

				(11.7 to 2018.7)			(0.65 to 0.99)	(0.76 to 0.99)		
Anismus	**EP**	12 (522)	21.7 (15.2 to 29.4)	290.6 (64.3 to 2641)	0.90 (0.73 to 0.99)	0.97 (0.93 to 0.99)	0.88 (0.76 to 0.96)	0.97 (0.91 to 1.00)	27.44 (13.16 to 91.43)	0.1 (0.02 to 0.28)
	MRI	6 (249)		162.4 (24.5 to 1913.5)	0.88 (0.60 to 0.99)	0.96 (0.89 to 0.98)	0.84 (0.65 to 0.95)	0.97 (0.88 to 1.00)	19.21 (7.58 to 55.55)	0.13 (0.01 to 0.42)
	TPUS	3 (226)		85.3 (12.4 to 657.2)	0.86 (0.58 to 0.97)	0.93 (0.81 to 0.98)	0.77 (0.51 to 0.92)	0.96 (0.87 to 0.99)	12.44 (4.18 to 35.76)	0.16 (0.03 to 0.46)
	EVUS	1 (131)		27.4 (3.1 to 457.7)	0.85 (0.53 to 0.99)	0.83 (0.56 to 0.95)	0.57 (0.30 to 0.83)	0.95 (0.85 to 1.00)	4.86 (1.73 to 15.53)	0.19 (0.02 to 0.6)
	DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	4 (169)		66.5 (10.8 to 642.9)	0.88 (0.73 to 0.97)	0.90 (0.68 to 0.98)	0.70 (0.39 to 0.94)	0.96 (0.91 to 0.99)	8.43 (2.63 to 52.65)	0.14 (0.04 to 0.33)
Pelvic floor descent	**EP**	7 (359)	53.5 (40 to 67.8)	167.6 (20.4 to 2049.1)	0.96 (0.86 to 0.99)	0.88 (0.63 to 0.98)	0.90 (0.70 to 0.99)	0.95 (0.79 to 0.99)	7.6 (2.54 to 43.71)	0.05 (0.01 to 0.18)
	MRI	4 (233)		35.6 (4.7 to 572.9)	0.88 (0.60 to 0.98)	0.83 (0.55 to 0.98)	0.85 (0.63 to 0.98)	0.85 (0.57 to 0.97)	4.9 (1.89 to 43.08)	0.15 (0.03 to 0.5)
	TPUS	1 (54)		134.7 (5.9 to 3629.4)	0.87 (0.55 to 0.99)	0.95 (0.62 to 1.00)	0.95 (0.69 to 1.00)	0.86 (0.59 to 0.99)	16.82 (2.11 to 175.98)	0.14 (0.01 to 0.5)

Table 18. Sensitivity analysis 2: DTA characteristics in women with ODS only *(Continued)*

EVUS	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	2 (99)	47.0 (4.8 to 730.2)	0.92 (0.62 to 0.99)	0.79 (0.56 to 0.97)	0.83 (0.63 to 0.98)	0.89 (0.59 to 0.99)	4.28 (1.93 to 26.47)	0.10 (0.01 to 0.49)
EDF	1 (29)	69.1 (4.3 to 1832.2)	0.84 (0.54 to 0.98)	0.93 (0.60 to 0.99)	0.93 (0.67 to 0.99)	0.83 (0.56 to 0.98)	10.8 (1.95 to 111.28)	0.18 (0.02 to 0.54)

Table 19. Sensitivity analysis 2: Probability that index test is equal or better than EP

Target condition	Index test	Pooled sensitivity			Pooled specificity		
		Estimate (%) (95% CrI)	Difference vs EP (%) (95% CrI)	Probability	Estimate (%) (95% CrI)	Difference vs EP (%)(95% CrI)	Probability
Rectocele	**EP**	96.9 (90.9 to 99.3)	N/A	N/A	81.9 (61.6 to 94.2)	N/A	N/A
	MRI	88.9 (66.4 to 97.6)	−7.7 (−30.6 to 2.9)	0.095	84.4 (66.4 to 96.0)	2.9 (−19.6 to 26.7)	0.597
	TPUS	79.3 (55.8 to 94.9)	−17.2 (−40.6 to −0.7)	0.018	84.8 (72.7 to 94.1)	3.1 (−14.8 to 25)	0.631
	EVUS	59.7 (50.3 to 87.8)	−36.6 (−47.4 to −8.4)	0.003	88.5 (57.2 to 97.4)	6.0 (−26.2 to 29)	0.687
	DAE	72.9 (53.1 to 90.9)	−23.7 (−44 to −4.6)	0.005	89.6 (62.7 to 98.6)	7.1 (−20.6 to 29.3)	0.743
	EDF	96.0 (85.4 to 99.3)	−0.8 (−11.5 to 5.8)	0.390	89.1 (60.4 to 98.6)	6.7 (−23.3 to 30.2)	0.701
Enterocele	**EP**	89.6 (81.2 to 96.2)	N/A	N/A	97.5 (94.4 to 99.4)	N/A	N/A
	MRI	77.6 (60.2 to 90.2)	−11.9 (−29.7 to 2.9)	0.056	97.2 (92.2 to 99.3)	−0.4 (−5.4 to 3.3)	0.416
	TPUS	80.9 (58.9 to 95.4)	−8.6 (−31.1 to 8.2)	0.179	97.7 (91.5 to 99.6)	0.2 (−6.2 to 3.9)	0.531
	EVUS	77.2 (53.5 to 95)	−12.4 (−36.6 to 7.3)	0.134	94.0 (68.2 to 98.7)	−3.5 (−29 to 2.3)	0.161
	DAE	74.3 (52.5 to 94.4)	−15 (−38.1 to 6.7)	0.106	96.8 (75.7 to 99.6)	−0.7 (−21.7 to 3.6)	0.398

Table 19. Sensitivity analysis 2: Probability that index test is equal or better than EP *(Continued)*

	EDF	68.3 (51.1 to 94.4)	−21 (−40.4 to 6.2)	0.085	97.6 (87.5 to 99.6)	0.0 (−10.2 to 3.9)	0.494
Intussusception	**EP**	87.5 (77.0 to 96.0)	N/A	N/A	91.9 (85.1 to 97.7)	N/A	N/A
	MRI	58.1 (50.4 to 77.7)	−28.7 (−41.7 to −8.1)	0.004	96.6 (87.7 to 99.5)	4.3 (−5.8 to 11.7)	0.834
	TPUS	75.1 (52.8 to 93.7)	−12.2 (−36.4 to 9)	0.151	95.0 (88.3 to 98.7)	2.9 (−5.8 to 10.9)	0.744
	EVUS	67.9 (51 to 94.8)	−19.2 (−39.5 to 8.7)	0.112	90.3 (59.3 to 98.5)	−1.9 (−33.4 to 9.2)	0.400
	DAE	61.5 (50.4 to 89.5)	−25.1 (−41.2 to 3.6)	0.048	92.7 (62.9 to 99.0)	0.2 (−29.5 to 10.5)	0.509
	EDF	89.2 (64.4 to 98.5)	1.3 (−24 to 16.4)	0.556	92.3 (71.5 to 99.0)	0 (−21 to 10.4)	0.502
Anismus	**EP**	90.0 (73.2 to 98.5)	N/A	N/A	96.7 (93.4 to 99.0)	N/A	N/A
	MRI	87.8 (59.9 to 98.7)	−2.3 (−29.1 to 16.2)	0.405	95.6 (89.4 to 98.4)	−1.2 (−7.5 to 3.1)	0.293
	TPUS	85.7 (57.8 to 97.3)	−4.3 (−32.6 to 15.4)	0.335	93.3 (80.9 to 97.7)	−3.5 (−15.8 to 2)	0.112
	EVUS	84.8 (53.4 to 98.6)	−4.9 (−38.2 to 17.7)	0.357	83.3 (56.0 to 94.6)	−13.3 (−40.8 to −1.7)	0.010
	DAE	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	87.8 (72.6 to 96.4)	−2.1 (−19.4 to 16)	0.395	89.6 (67.6 to 98.3)	−7 (−28.9 to 2.3)	0.097
PFD	**EP**	95.8 (85.7 to 99.2)	N/A	N/A	87.5 (63.1 to 97.8)	N/A	N/A
	MRI	87.8 (59.5 to 97.5)	−7.6 (−35.9 to 5.8)	0.137	82.9 (54.7 to 98.1)	−4 (−34.8 to 24)	0.382
	TPUS	87.2 (55.4 to 98.7)	−8 (−40.1 to 7.1)	0.183	95.0 (62.0 to 99.5)	6 (−26.2 to 31.3)	0.730
	EVUS	N/A	N/A	N/A	N/A	N/A	N/A
	DAE	91.9 (62.1 to 99.1)	−3.5 (−33.2 to 8.2)	0.283	79.1 (55.5 to 96.7)	−7 (−33.4 to 19.9)	0.296
	EDF	83.8 (54.2 to 98.3)	−11.3 (−41.2 to 5.7)	0.129	92.5 (60.1 to 99.3)	4 (−27.6 to 29.6)	0.651

Probability of < 0.400 means estimated test accuracy of index test is not equal or higher than EP
Probability of > 0.400 means estimated test accuracy of index test is equal or higher than EP

Table 20. Sensitivity analysis 3: DTA characteristics for current methodology (studies published after 2009)

Target condition	Imaging	Number of studies (partici-pants)	Preva-lence (%)(95% CI)	Diagnostic Odds Ratio (95% CrI)	Sensitivity (95% CrI)	Specificity (95% CrI)	PPV (95% CrI)	NPV (95% CrI)	LR+ (95% CrI)	LR−(95% CrI)
Rectocele	**EP**	20 (1379)	65.1 (57.4 - 72.7)	132.9 (40.2 to 591.3)	0.97 (0.91 to 0.99)	0.81 (0.72 to 0.89)	0.91 (0.84 to 0.95)	0.93 (0.81 to 0.98)	5.10 (3.48 to 8.80)	0.04 (0.01 to 0.11)
	MRI	11 (522)		54.3 (11 to 340.9)	0.91 (0.81 to 0.97)	0.83 (0.62 to 0.96)	0.91 (0.80 to 0.98)	0.83 (0.66 to 0.95)	5.34 (2.33 to 21.15)	0.11 (0.04 to 0.25)
	TPUS	7 (840)		119.5 (24.8 to 849.5)	0.93 (0.80 to 0.99)	0.89 (0.78 to 0.96)	0.94 (0.88 to 0.98)	0.88 (0.68 to 0.98)	8.38 (4.15 to 25.33)	0.07 (0.01 to 0.23)
	EVUS	2 (454)		8.5 (2.3 to 40.5)	0.71 (0.52 to 0.89)	0.77 (0.54 to 0.93)	0.85 (0.72 to 0.95)	0.58 (0.41 to 0.79)	3.08 (1.51 to 10.01)	0.39 (0.15 to 0.68)
	DAE	1 (56)		20.9 (2.7 to 321.5)	0.76 (0.53 to 0.93)	0.86 (0.55 to 0.99)	0.91 (0.73 to 0.99)	0.65 (0.44 to 0.87)	5.49 (1.57 to 58.41)	0.29 (0.09 to 0.63)
	EDF	3 (139)		127.1 (11.3 to 2264.2)	0.94 (0.78 to 0.99)	0.88 (0.57 to 0.99)	0.94 (0.78 to 0.99)	0.89 (0.65 to 0.98)	7.71 (2.10 to 75.81)	0.07 (0.01 to 0.27)
Entero-cele	**EP**	19 (1932)	17.5 (13.9 to 21.5)	320.9 (107.7 to 1500.6)	0.93 (0.86 to 0.98)	0.96 (0.93 to 0.99)	0.82 (0.70 to 0.94)	0.98 (0.97 to 1.00)	22.06 (12.60 to 61.74)	0.07 (0.02 to 0.15)
	MRI	10 (1075)		515.6 (98.6 to 3650.4)	0.79 (0.65 to 0.91)	0.99 (0.97 to 1.00)	0.96 (0.83 to 0.99)	0.96 (0.92 to 0.98)	104.72 (24.29 to 586.55)	0.21 (0.09 to 0.36)

Table 20. Sensitivity analysis 3: DTA characteristics for current methodology (studies published after 2009) *(Continued)*

	TPUS	7 (840)		213.9 (33.3 to 2015.3)	0.80 (0.57 to 0.95)	0.98 (0.93 to 1.00)	0.89 (0.69 to 0.98)	0.96 (0.91 to 0.99)	39.43 (11.07 to 217.22)	0.20 (0.05 to 0.45)
	EVUS	2 (454)		57.8 (5.8 to 336.4)	0.60 (0.50 to 0.86)	0.97 (0.78 to 0.99)	0.83 (0.37 to 0.95)	0.92 (0.88 to 0.97)	22.64 (2.85 to 95.11)	0.42 (0.15 to 0.54)
	DAE	1 (56)		42.1 (3.0 to 590.4)	0.66 (0.51 to 0.90)	0.95 (0.62 to 1.00)	0.75 (0.27 to 0.97)	0.93 (0.88 to 0.98)	13.93 (1.76 to 136.62)	0.37 (0.10 to 0.62)
	EDF	3 (139)		108.7 (14.2 to 1697.7)	0.72 (0.52 to 0.96)	0.97 (0.88 to 1.00)	0.86 (0.54 to 0.97)	0.94 (0.89 to 0.99)	28.11 (5.66 to 159.12)	0.29 (0.04 to 0.50)
Intussus-ception	**EP**	20 (1384)	46.9 (36.2 to 55.8)	82.2 (26.4 to 451)	0.87 (0.76 to 0.96)	0.92 (0.85 to 0.98)	0.91 (0.80 to 0.98)	0.89 (0.78 to 0.97)	10.88 (5.79 to 38.86)	0.14 (0.05 to 0.27)
	MRI	11 (527)		51.6 (10.2 to 439.8)	0.62 (0.51 to 0.80)	0.97 (0.87 to 1.00)	0.95 (0.77 to 0.99)	0.74 (0.64 to 0.87)	19.61 (4.77 to 140.46)	0.40 (0.21 to 0.52)
	TPUS	6 (517)		48.1 (12.8 to 361)	0.67 (0.51 to 0.90)	0.96 (0.88 to 0.99)	0.93 (0.81 to 0.99)	0.77 (0.64 to 0.92)	15.79 (5.71 to 71.77)	0.35 (0.11 to 0.51)
	EVUS	2 (454)		23.8 (4.5 to 188.8)	0.63 (0.51 to 0.86)	0.93 (0.73 to 0.99)	0.89 (0.66 to 0.98)	0.74 (0.62 to 0.89)	9.11 (2.35 to 54.37)	0.41 (0.16 to 0.56)
	DAE	1 (56)		18.8 (1.8 to 344)	0.62 (0.51 to 0.91)	0.91 (0.55 to 0.99)	0.86 (0.53 to 0.99)	0.73 (0.58 to 0.92)	6.89 (1.36 to 83.85)	0.43 (0.11 to 0.75)
	EDF	3 (139)		94.6	0.93	0.88	0.87	0.93	7.44	0.09

Table 20. Sensitivity analysis 3: DTA characteristics for current methodology (studies published after 2009) *(Continued)*

				(9.9 to 1280.2)	(0.74 to 0.99)	(0.63 to 0.98)	(0.64 to 0.98)	(0.77 to 0.99)	(2.36 to 44.37)	(0.01 to 0.33)
Anismus	**EP**	12 (904)	23.2 (16.3 to 31.1)	68.3 (27.7 to 253.3)	0.68 (0.54 to 0.86)	0.97 (0.94 to 0.99)	0.87 (0.75 to 0.95)	0.91 (0.84 to 0.97)	21.93 (11.37 to 58.05)	0.33 (0.15 to 0.48)
	MRI	6 (277)		113.9 (20.8 to 1213.8)	0.80 (0.55 to 0.97)	0.96 (0.90 to 0.99)	0.87 (0.66 to 0.96)	0.94 (0.85 to 0.99)	21.36 (7.44 to 75.3)	0.21 (0.03 to 0.47)
	TPUS	4 (610)		89.5 (15.6 to 689.5)	0.90 (0.64 to 0.98)	0.91 (0.80 to 0.97)	0.75 (0.54 to 0.90)	0.97 (0.88 to 0.99)	9.74 (4.32 to 26.65)	0.11 (0.02 to 0.40)
	EVUS	2 (454)		57.9 (6.0 to 428.7)	0.85 (0.59 to 0.96)	0.91 (0.64 to 0.98)	0.73 (0.38 to 0.92)	0.95 (0.87 to 0.99)	9.26 (2.18 to 33.95)	0.17 (0.04 to 0.48)
	DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	3 (139)		63.2 (8.1 to 842.3)	0.85 (0.65 to 0.96)	0.91 (0.68 to 0.99)	0.74 (0.40 to 0.96)	0.95 (0.88 to 0.99)	9.59 (2.48 to 78.58)	0.17 (0.04 to 0.41)
Pelvic floor descent	**EP**	8 (398)	64 (49.5 to 77.9)	127.8 (19.5 to 1259.4)	0.97 (0.89 to 0.99)	0.80 (0.56 to 0.95)	0.90 (0.74 to 0.98)	0.93 (0.73 to 0.99)	4.86 (2.16 to 20.54)	0.04 (0.01 to 0.15)
	MRI	6 (315)		72.0 (10.1 to 1122)	0.95 (0.79 to 0.99)	0.80 (0.53 to 0.98)	0.89 (0.72 to 0.99)	0.89 (0.61 to 0.98)	4.75 (1.99 to 42.86)	0.07 (0.02 to 0.27)
	TPUS	1 (54)		144.4 (6.1 to 4013.6)	0.88 (0.56 to 0.99)	0.95 (0.62 to 1.00)	0.97 (0.77 to 1.00)	0.81 (0.48 to 0.98)	17.07 (2.11 to 175.66)	0.13 (0.01 to 0.50)

Table 20. Sensitivity analysis 3: DTA characteristics for current methodology (studies published after 2009) *(Continued)*

EVUS	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	1 (56)	12.0 (2.1 to 165.1)	0.81 (0.54 to 0.97)	0.71 (0.52 to 0.95)	0.84 (0.66 to 0.97)	0.68 (0.39 to 0.94)	2.77 (1.44 to 15.55)	0.27 (0.04 to 0.69)
EDF	1 (29)	73.5 (4.8 to 1956.4)	0.84 (0.54 to 0.98)	0.93 (0.60 to 0.99)	0.95 (0.76 to 1.00)	0.76 (0.45 to 0.97)	11.1 (1.97 to 111.82)	0.18 (0.02 to 0.53)

Table 21. Sensitivity analysis 3: Probability that index test is equal or better than EP

Target condition	Index test	Pooled sensitivity			Pooled specificity		
		Estimate (%) (95% CrI)	Difference vs EP (%)	Probability	Estimate (%) (95% CrI)	Difference vs EP (%)	Probability
Rectocele	**EP**	96.8 (91.1 to 99.2)	N/A	N/A	81.1 (72.4 to 89.1)	N/A	N/A
	MRI	91.3 (80.9 to 97.1)	−5.2 (−16.0 to 2.5)	0.009	83.0 (62.2 to 95.7)	1.8 (−19.8 to 16.9)	0.576
	TPUS	93.4 (80.4 to 98.8)	−3.3 (−16.5 to 4.4)	0.212	88.9 (78.3 to 96.3)	7.7 (−5.3 to 19.3)	0.899
	EVUS	70.6 (52.0 to 88.6)	−25.8 (−44.7 to −7.3)	0.002	77.4 (53.7 to 93.4)	3.9 (−28.5 to 14.5)	0.360
	DAE	75.9 (53.1 to 92.6)	−20.5 (−43.5 to −3.2)	0.007	86.4 (54.7 to 98.7)	5.1 (−27.1 to 20.6)	0.639
	EDF	94.3 (78.1 to 98.8)	−2.4 (−18.7 to 4.8)	0.247	88.0 (56.5 to 98.8)	6.5 (−25.1 to 21.2)	0.683
Enterocele	**EP**	93.0 (85.5 to 97.8)	N/A	N/A	95.8 (92.7 to 98.5)	N/A	N/A
	MRI	79.1 (64.7 to 91.0)	−13.6 (−28.7 to 0.1)	0.026	99.3 (96.7 to 99.9)	3.3 (−0.3 to 6.6)	0.968
	TPUS	80.0 (56.8 to 95.3)	−12.8 (−36.4 to 4.2)	0.086	98.0 (93.4 to 99.6)	2.0 (−3.1 to 5.9)	0.831
	EVUS	59.9 (50.4 to 85.7)	−32.6 (−44.6 to −6.7)	0.005	97.3 (77.9 to 99.4)	1.2 (−18.1 to 5.4)	0.642
	DAE	65.5 (50.8 to 90.4)	−26.9 (−43.6 to −1.6)	0.017	95.2 (61.9 to 99.5)	−0.8 (−33.8 to 5.2)	0.434

Table 21. Sensitivity analysis 3: Probability that index test is equal or better than EP *(Continued)*

	EDF	72.0 (51.6 to 96.4)	−20.8 (−42.1 to 4.9)	0.079	97.4 (87.7 to 99.5)	1.4 (−8.2 to 5.6)	0.709
Intussusception	**EP**	86.9 (75.6 to 95.7)	N/A	N/A	92.0 (85.2 to 97.8)	N/A	N/A
	MRI	61.9 (51 to 80.1)	−24.6 (−39.5 to −4.1)	0.010	96.8 (87.3 to 99.6)	4.2 (−5.9 to 11.8)	0.832
	TPUS	66.6 (51.1 to 89.5)	−20 (−38.9 to 5)	0.063	95.8 (88.4 to 99)	3.5 (−5.5 to 11.2)	0.791
	EVUS	62.5 (51.2 to 85.7)	−23.8 (−39.3 to 0.1)	0.026	93.0 (73.1 to 98.8)	0.8 (−19.3 to 9.5)	0.560
	DAE	62.4 (50.5 to 90.6)	−23.7 (−40.7 to 5.7)	0.063	90.7 (55.2 to 99.2)	−1.8 (−36.7 to 10.5)	0.433
	EDF	92.6 (73.5 to 98.8)	5.1 (−14.7 to 18.8)	0.742	87.8 (63.2 to 97.9)	−4.6 (−29.6 to 7.8)	0.265
Anismus	**EP**	67.6 (53.6 to 85.8)	N/A	N/A	96.9 (94.2 to 98.8)	N/A	N/A
	MRI	80.2 (55.2 to 97.1)	11.8 (−15 to 33.2)	0.807	96.3 (90.1 to 98.9)	−0.6 (−6.8 to 3.2)	0.388
	TPUS	89.6 (63.6 to 98.2)	20.5 (−7.8 to 38.8)	0.934	91.0 (80.0 to 96.7)	−5.8 (−16.8 to 0.3)	0.030
	EVUS	84.8 (59.3 to 96.4)	16.1 (−13.3 to 36.1)	0.877	91.0 (63.7 to 97.6)	−5.8 (−33.2 to 1.1)	0.067
	DAE	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	85.1 (65.4 to 95.9)	16.6 (−7.5 to 35.2)	0.923	91.2 (67.7 to 98.9)	−5.6 (−29.4 to 2.6)	0.148
PFD	**EP**	96.8 (88.9 to 99.4)	N/A	N/A	80.2 (56 to 95.3)	N/A	N/A
	MRI	94.5 (79.0 to 98.7)	−2.1 (−17.7 to 6.4)	0.275	80.4 (53.1 to 97.8)	0.4 (−31.3 to 30.6)	0.508
	TPUS	87.8 (55.8 to 98.7)	−8.6 (−40.4 to 4.6)	0.138	95.1 (61.6 to 99.5)	12.7 (−20.4 to 38.8)	0.841
	EVUS	N/A	N/A	N/A	N/A	N/A	N/A
	DAE	80.8 (54.0 to 97.1)	−15.4 (−42.3 to 1.9)	0.049	71.3 (51.7 to 95.0)	−7.5 (−35 to 24.2)	0.324
	EDF	84.3 (54.3 to 98.4)	−11.9 (−42 to 3.5)	0.096	92.7 (59.9 to 99.3)	10.6 (−22.5 to 37.6)	0.792

Probability of < 0.400 means estimated test accuracy of index test is not equal or higher than EP
Probability of > 0.400 means estimated test accuracy of index test is equal or higher than EP

Table 22. Sensitivity analysis 4: DTA characteristics high level of evidence (high risk of bias excluded, only women with ODS and current methodology of tests)

Target condition	Imaging	Number of studies (patients)	Prevalence (%)(95% CI)	Diagnostic Odds Ratio (95% CI)	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	LR+ (95% CI)	LR- (95% CI)
Rectocele	**EP**	11 (610)	56.4 (45.8 to 68)	102.9 (21 to 1043.7)	0.95 (0.82 to 0.99)	0.84 (0.74 to 0.96)	0.88 (0.78 to 0.98)	0.92 (0.75 to 0.99)	5.84 (3.57 to 25.06)	0.06 (0.01 to 0.21)
	MRI	6 (348)		77.0 (12.4 to 645.2)	0.93 (0.78 to 0.98.3)	0.85 (0.64 to 0.97)	0.89 (0.74 to 0.98)	0.90 (0.69 to 0.98)	6.21 (2.5 to 30.77)	0.09 (0.02 to 0.28)
	TPUS	4 (274)		31.1 (5.6 to 313.3)	0.86 (0.60 to 0.98)	0.83 (0.68 to 0.93)	0.87 (0.74 to 0.95)	0.82 (0.54 to 0.97)	4.95 (2.41 to 12.6)	0.17 (0.02 to 0.50)
	EVUS	1 (131)		13.2 (2.1 to 99.6)	0.60 (0.51 to 0.88)	0.89 (0.58 to 0.97)	0.88 (0.64 to 0.97)	0.63 (0.47 to 0.85)	5.51 (1.46 to 24.5)	0.46 (0.15 to 0.71)
	DAE	1 (56)		21.6 (2.7 to 350.9)	0.76 (0.53 to 0.93)	0.87 (0.55 to 0.99)	0.88 (0.65 to 0.99)	0.73 (0.51 to 0.91)	5.65 (1.57 to 60.3)	0.29 (0.09 to 0.63)
	EDF	2 (110)		53.5 (5.1 to 999.3)	0.92 (0.66 to 0.98)	0.82 (0.53 to 0.98)	0.87 (0.67 to 0.99)	0.88 (0.61 to 0.97)	5.01 (1.86 to 54.31)	0.10 (0.02 to 0.45)
Entero-cele	**EP**	10 (557)	22.5 (16.2 to 29)	283.2 (66.1 to 1687.6)	0.90 (0.79 to 0.97)	0.97 (0.92 to 0.99)	0.89 (0.73 to 0.97)	0.97 (0.93 to 0.99)	26.63 (10.8 to 106.58)	0.10 (0.03 to 0.22)
	MRI	5 (295)		100.5 (24.3 to 548.3)	0.75 (0.59 to 0.88)	0.97 (0.91 to 0.99)	0.88 (0.68 to 0.97)	0.93 (0.88 to 0.97)	24.91 (8.17 to 107.83)	0.26 (0.12 to 0.42)

Table 22. Sensitivity analysis 4: DTA characteristics high level of evidence (high risk of bias excluded, only women with ODS and current methodology of tests) *(Continued)*

	TPUS	4 (274)		111.0 (11.4 to 1246.2)	0.74 (0.53 to 0.94)	0.97 (0.83 to 1.00)	0.89 (0.53 to 0.98)	0.93 (0.86 to 0.98)	27.27 (4.19 to 188.4)	0.27 (0.06 to 0.50)
	EVUS	1 (131)		35.6 (3.2 to 331)	0.68 (0.51 to 0.92)	0.94 (0.62 to 0.99)	0.77 (0.33 to 0.95)	0.91 (0.83 to 0.98)	11.48 (1.79 to 58.47)	0.35 (0.09 to 0.60)
	DAE	1 (56)		42.8 (3.1 to 618.9)	0.66 (0.51 to 0.90)	0.95 (0.63 to 1.00)	0.80 (0.33 to 0.98)	0.91 (0.83 to 0.97)	13.99 (1.77 to 136.95)	0.37 (0.11 to 0.61)
	EDF	2 (110)		86.0 (7.1 to 1356.2)	0.67 (0.51 to 0.95)	0.97 (0.78 to 1.00)	0.88 (0.46 to 0.98)	0.91 (0.85 to 0.98)	25.76 (3.09 to 204.69)	0.34 (0.06 to 0.52)
Intussus-ception	**EP**	11 (610)	43.1 (30.8 to 54)	86.8 (20 to 541.7)	0.88 (0.77 to 0.97)	0.97 (0.86 to 1.00)	0.89 (0.71 to 0.97)	0.91 (0.81 to 0.98)	10.29 (4.35 to 37.68)	0.13 (0.04 to 0.27)
	MRI	6 (348)		45.3 (8.4 to 356.6)	0.60 (0.51 to 0.81)	0.95 (0.88 to 0.99)	0.93 (0.72 to 0.99)	0.76 (0.65 to 0.9)	17.79 (4.17 to 113.05)	0.42 (0.19 to 0.53)
	TPUS	4 (274)		58.3 (12.8 to 496.6)	0.74 (0.52 to 0.94)	0.90 (0.59 to 0.98)	0.92 (0.78 to 0.98)	0.83 (0.68 to 0.96)	14.95 (5.51 to 63.25)	0.28 (0.07 to 0.51)
	EVUS	1 (131)		22.4 (2.7 to 280.3)	0.69 (0.51 to 0.95)	0.91 (0.56 to 0.99)	0.84 (0.52 to 0.97)	0.79 (0.64 to 0.96)	6.78 (1.65 to 44.1)	0.35 (0.06 to 0.64)
	DAE	1 (56)		19.0 (1.9 to 376.5)	0.62 (0.51 to 0.91)	0.91 (0.63 to 0.99)	0.84 (0.49 to 0.99)	0.76 (0.6 to 0.93)	6.93 (1.38 to 87.51)	0.43 (0.10 to 0.73)

Table 22. Sensitivity analysis 4: DTA characteristics high level of evidence (high risk of bias excluded, only women with ODS and current methodology of tests) *(Continued)*

	EDF	2 (110)		95.1 (7.6 to 1503.9)	0.90 (0.64 to 0.99)	0.91 (0.81 to 0.98)	0.88 (0.61 to 0.98)	0.92 (0.74 to 0.99)	9.38 (2.24 to 68.56)	0.11 (0.02 to 0.44)
Anismus	**EP**	6 (362)	19.4 (11.4 to 30.6)	69.8 (18.3 to 550.5)	0.76 (0.54 to 0.95)	0.96 (0.90 to 0.99)	0.80 (0.58 to 0.94)	0.94 (0.85 to 0.99)	16.71 (7.04 to 53.51)	0.26 (0.05 to 0.48)
	MRI	3 (189)		71.3 (10.7 to 945.1)	0.76 (0.52 to 0.97)	0.96 (0.83 to 0.99)	0.80 (0.47 to 0.95)	0.94 (0.84 to 0.99)	16.87 (4.13 to 60.83)	0.25 (0.03 to 0.51)
	TPUS	2 (185)		33.8 (5.3 to 358.8)	0.73 (0.52 to 0.96)	0.93 (0.72 to 0.98)	0.69 (0.34 to 0.91)	0.94 (0.84 to 0.99)	9.53 (2.44 to 32.97)	0.30 (0.05 to 0.55)
	EVUS	1 (131)		30.1 (3.2 to 497)	0.85 (0.54 to 0.99)	0.81 (0.56 to 0.95)	0.55 (0.26 to 0.83)	0.96 (0.85 to 1)	5.1 (1.75 to 16.88)	0.18 (0.02 to 0.59)
	DAE	0 (0)		N/A	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	1 (110)		40.3 (4.4 to 575.3)	0.85 (0.59 to 0.97)	0.87 (0.58 to 0.98)	0.61 (0.27 to 0.94)	0.96 (0.87 to 0.99)	6.41 (1.9 to 52.98)	0.18 (0.04 to 0.51)
Pelvic floor descent	**EP**	4 (249)	52.6 (33.9 to 69.1)	82.7 (11.7 to 1134.7)	0.95 (0.84 to 0.99)	0.80 (0.54 to 0.97)	0.83 (0.59 to 0.98)	0.94 (0.77 to 0.99)	4.74 (2.03 to 34.3)	0.06 (0.02 to 0.23)
	MRI	3 (195)		64.7 (7.1 to 1204.9)	0.82 (0.53 to 0.97)	0.93 (0.71 to 0.99)	0.93 (0.70 to 0.99)	0.82 (0.53 to 0.98)	11.32 (2.64 to 100.07)	0.20 (0.03 to 0.52)
	TPUS	1 (54)		122.2 (5.3 to 3402.5)	0.87 (0.56 to 0.99)	0.95 (0.63 to 1.00)	0.95 (0.67 to 1)	0.87	15.36 (2.1 to 165.07)	0.14

Table 22. Sensitivity analysis 4: DTA characteristics high level of evidence (high risk of bias excluded, only women with ODS and current methodology of tests) *(Continued)*

						(0.58 to 0.98)		(0.02 to 0.50)
EVUS	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DAE	1 (56)	14.3 (2.4 to 212.1)	0.78 (0.53 to 0.96)	0.78 (0.53 to 0.97)	0.80 (0.52 to 0.97)	0.77 (0.48 to 0.96)	3.47 (1.52 to 24.54)	0.29 (0.05 to 0.66)
EDF	0 (0)	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 23. Sensitivity analysis 4: Probability index test is equal or better than EP

Target condition	Index test	Pooled sensitivity			Pooled specificity		
		Estimate (%) (95% CrI)	Difference vs EP (%)	Probability	Estimate (95% CrI)	Difference vs EP (%)	Probability
Rectocele	**EP**	94.7 (82.1 to 99)	N/A	N/A	83.9 (74.3 to 96.3)	N/A	N/A
	MRI	92.5 (77.7 to 98.3)	−2 (−17 to 11.4)	0.356	85.2 (64.2 to 97.1)	0.4 (−21.3 to 16)	0.517
	TPUS	85.8 (59.6 to 98)	−8.3 (−34.8 to 8.8)	0.180	83.1 (68 to 93.2)	−1.4 (−19.5 to 12.6)	0.427
	EVUS	60 (50.5 to 87.6)	−33.4 (−46.2 to −4.6)	0.009	88.9 (57.9 to 97.4)	3.8 (−27.8 to 18.1)	0.637
	DAE	75.7 (53.4 to 92.6)	−18.2 (−41.6 to 1.6)	0.037	87 (55.2 to 98.7)	1.7 (−30.6 to 18.9)	0.547
	EDF	91.8 (66.3 to 98.1)	−2.8 (−28 to 10.9)	0.314	82.1 (53.1 to 98.4)	−2.9 (−33.4 to 17.9)	0.425
Enterocele	**EP**	90.2 (78.8 to 97)	N/A	N/A	96.6 (91.9 to 99.1)	N/A	N/A
	MRI	75.2 (59.3 to 88.4)	−14.7 (−32.2 to 2.2)	0.044	97 (91 to 99.3)	0.3 (−5.9 to 5.5)	0.550
	TPUS	74.2 (52.7 to 93.9)	−15.4 (−38.4 to 6.5)	0.097	97.3 (83.3 to 99.6)	0.4 (−13.3 to 6)	0.563
	EVUS	67.7 (51.1 to 91.6)	−21.7 (−40.8 to 3.6)	0.052	94.2 (62.3 to 98.8)	−2.5 (−34.2 to 4.3)	0.272
	DAE	65.9 (50.9 to 90.2)	−23.5 (−41.7 to 2.7)	0.040	95.3 (62.7 to 99.5)	−1.4 (−34 to 5.2)	0.373

Table 23. Sensitivity analysis 4: Probability index test is equal or better than EP *(Continued)*

	EDF	67.4 (50.9 to 94.5)	−22 (−41.6 to 6.1)	0.078	97.4 (78.1 to 99.7)	0.5 (−18.5 to 6)	0.566
Intussusception	**EP**	88.3 (76.5 to 96.6)	N/A	N/A	96.6 (85.9 to 99.5)	N/A	N/A
	MRI	59.8 (50.5 to 81.3)	−27.5 (−42.3 to −3.4)	0.013	95.2 (87.7 to 98.8)	4.7 (−6.6 to 15.5)	0.845
	TPUS	73.4 (51.8 to 93.5)	−14.8 (−38.6 to 8.2)	0.132	89.8 (58.8 to 98.4)	3.6 (−5.8 to 14.6)	0.788
	EVUS	69 (51 to 95)	−18.4 (−40.4 to 8.6)	0.121	90.7 (55.6 to 99.2)	−1.6 (−32.8 to 11.6)	0.415
	DAE	62.3 (50.5 to 90.9)	−24.9 (−42.4 to 5.4)	0.057	90.6 (62.9 to 98.7)	−0.9 (−36.5 to 14)	0.462
	EDF	90.2 (63.9 to 98.5)	1.4 (−24.8 to 16.3)	0.568	91.4 (80.5 to 97.6)	−1.1 (−28.7 to 12.5)	0.449
Anismus	**EP**	75.6 (54 to 95.1)	N/A	N/A	95.5 (89.6 to 98.6)	N/A	N/A
	MRI	75.9 (51.7 to 97)	0.2 (−30.2 to 28.5)	0.506	95.6 (82.7 to 98.7)	0 (−12.7 to 6.5)	0.498
	TPUS	72.8 (51.5 to 95.5)	−2.7 (−33.7 to 29.2)	0.434	92.5 (71.8 to 97.7)	−3 (−23.3 to 4.6)	0.211
	EVUS	85.3 (53.6 to 98.6)	7.9 (−28.9 to 36.5)	0.669	84.1 (56 to 95)	−11.3 (−39.2 to 0.7)	0.035
	DAE	N/A	N/A	N/A	N/A	N/A	N/A
	EDF	84.7 (59.1 to 96.8)	7.9 (−21.2 to 33.6)	0.713	87.1 (57.8 to 98.4)	−8.2 (−37.6 to 4.4)	0.143
PFD	**EP**	95.2 (83.6 to 98.7)	N/A	N/A	80.2 (54.3 to 97.3)	N/A	N/A
	MRI	82 (53 to 97.2)	−12.3 (−41.8 to 5.9)	0.120	93.1 (71.2 to 99.2)	11.6 (−14.6 to 39.5)	0.815
	TPUS	86.8 (55.6 to 98.6)	−7.7 (−39 to 7.5)	0.194	94.7 (63.1 to 99.5)	12.1 (−19.2 to 41.5)	0.826
	EVUS	N/A	N/A	N/A	N/A	N/A	N/A
	DAE	77.8 (53.3 to 95.9)	−16.6 (−42.1 to 3.6)	0.063	78.1 (53.4 to 96.8)	−1.3 (−32.6 to 30.5)	0.463
	EDF	N/A	N/A	N/A	N/A	N/A	N/A

Probability of < 0.400 means estimated test accuracy of index test is not equal or higher than EP
Probability of > 0.400 means estimated test accuracy of index test is equal or higher than EP

APPENDICES

Appendix 1. Search strategy for MEDLINE

Ovid MEDLINE 1950 to 18.12.2019

1. Defecography/
2. Fluoroscopy/
3. ra.fs.
4. (EP or defecogra* or defaecogra* or proctogra* or colpocystodefecogr* or colpocystorectogra* or videoproctogra* or cystocolpoproctogra* or fluorosco*).ti,ab,kw.
5. 1 or 2 or 3 or 4
6. Magnetic Resonance Imaging/
7. (Magnetic adj1 Resonance).ti,ab,kw.
8. MRI.ti,ab,kw.
9. (MR adj3 imaging).ti,ab,kw.
10. (dynamic adj3 MR*).ti,ab,kw.
11. 6 or 7 or 8 or 9 or 10
12. Ultrasonography/
13. Endosonography/
14. Imaging, Three-Dimensional/
15. us.fs.
16. (TPUS or EVUS or DEA or EFD or ultraso* or endosonogra* or echogra* or sonogra* or echodefecogra*).ti,ab,kw.
17. 12 or 13 or 14 or 15 or 16
18. Rectocele/
19. Intussusception/
20. Rectal Prolapse/
21. Rectal Diseases/
22. Rectum/
23. (rectocele or enterocele or sigmoidocele or intussusception or (rectal adj1 prolapse) or anismus or (perineal adj1 descent) or (pelvic adj1 floor adj1 descent) or (posterior adj1 compartment)).ti,ab,kw.
24. 18 or 19 or 20 or 21 or 22 or 23
25. ODS.ti,ab,kw.
26. (obstruct* adj1 (defecat* or defaecat*)).ti,ab,kw.
27. ((defecat* or defaecat* or evacuat* or anorect*) adj3 (disorder* or dysfunct* or difficult*)).ti,ab,kw.
28. 25 or 26 or 27
29. 11 or 17
30. 24 or 28
31. 5 and 29 and 30

Appendix 2. Search strategy for Embase

Ovid Embase 1974 to 18.12.2019

1. defecography/
2. fluoroscopy/
3. radiography/
4. (EP or defecogra* or defaecogra* or proctogra* or colpocystodefecogra* or colpocystorectogra* or videoproctogra* or cystocolpoproctogra* or fluorosco*).ti,ab,kw.
5. 1 or 2 or 3 or 4
6. nuclear magnetic resonance imaging/
7. (Magnetic adj1 Resonance).ti,ab,kw.
8. MRI.ti,ab,kw.
9. (MR adj3 imaging).ti,ab,kw.
10. (dynamic adj3 MR*).ti,ab,kw.
11. 6 or 7 or 8 or 9 or 10
12. echography/
13. transvaginal echography/
14. transrectal ultrasonography/
15. (TPUS or EVUS or DEA or EFD or ultraso* or endosonogra* or echogra* or sonogra* or echodefecogra*).ti,ab,kw.
16. 12 or 13 or 14 or 15
17. rectocele/
18. intussusception/
19. enterocele/

20. rectum prolapse/
21. rectum disease/
22. (rectocele or enterocele or sigmoidocele or intussusception or (rectal adj1 prolapse) or anismus or (perineal adj1 descent) or (pelvic adj1 floor adj1 descent) or (posterior adj1 compartment)).ti,ab,kw.
23. 17 or 18 or 19 or 20 or 21 or 22
24. defecation disorder/
25. ODS.ti,ab,kw.
26. (obstruct* adj1 (defecat* or defaecat*)).ti,ab,kw.
27. ((defecat* or defaecat* or evacuat* or anorect*) adj3 (disorder* or dysfunct* or difficult*)).ti,ab,kw.
28. 24 or 25 or 26 or 27
29. 11 or 16
30. 23 or 28
31. 5 and 29 and 30

Appendix 3. Search strategy for Cochrane Library

Cochrane Library searched 18.12.2019
#1 MeSH descriptor: [Defecography] explode all trees
#2 MeSH descriptor: [Fluoroscopy] explode all trees
#3 (EP or defecogra* or defaecogra* or proctogra* or colpocystodefecogr* or colpocystorectogra* or videoproctogra* or cystocolpoproctogra* or fluorosco*):ti,ab,kw
#4 (#1 or #2 or #3)
#5 MeSH descriptor: [Magnetic Resonance Imaging] explode all trees
#6 (Magnetic near/1 Resonance):ti,ab,kw
#7 MRI:ti,ab,kw
#8 (MR near/3 imaging):ti,ab,kw
#9 (dynamic near/3 MR*):ti,ab,kw
#10 (#5 or #6 or #7 or #8 or #9)
#11 MeSH descriptor: [Ultrasonography] explode all trees
#12 MeSH descriptor: [Endosonography] explode all trees
#13 MeSH descriptor: [Imaging, Three-Dimensional] explode all trees
#14 (TPUS or EVUS or DEA or EFD or ultraso* or endosonogra* or echogra* or sonogra* or echodefecogra*):ti,ab,kw
#15 (#11 or #12 or #13 or #14)
#16 MeSH descriptor: [Rectocele] explode all trees
#17 MeSH descriptor: [Intussusception] explode all trees
#18 MeSH descriptor: [Rectal Prolapse] explode all trees
#19 MeSH descriptor: [Rectal Diseases] explode all trees
#20 MeSH descriptor: [Rectum] explode all trees
#21 (rectocele or enterocele or sigmoidocele or intussusception or (rectal adj1 prolapse) or anismus or (perineal adj1 descent) or (pelvic adj1 floor adj1 descent) or (posterior adj1 compartment)):ti,ab,kw
#22 (#16 or #17 or #18 or #19 or #20 or #21)
#23 ODS:ti,ab,kw (Word variations have been searched) 58
#24 (obstruct* near/1 (defecat* or defaecat*)):ti,ab,kw
#25 (defecat* or defaecat* or evacuat* or anorect*) near/3 (disorder* or dysfunct* or difficult*):ti,ab,kw
#26 (#23 or #24 or #25)
#27 (#10 or #15)
#28 (#22 or #26)
#29 (#4 and #27 and #28)

Appendix 4. Search strategy for CINAHL

EBSCO CINAHL 1981 to 18.12.2019 1. Defecography/
2. Fluoroscopy/
3. Radiography/
4. (EP or defecogra* or defaecogra* or proctogra* or colpocystodefecogr* or colpocystorectogra* or videoproctogra* or cystocolpoproctogra* or fluorosco*).ti,ab
5. 1 or 2 or 3 or 4
6. Magnetic Resonance Imaging/
7. (Magnetic n1 Resonance).ti,ab
8. MRI.ti,ab
9. (MR n3 imaging).ti,ab
10. (dynamic n3 MR*).ti,ab
11. 6 or 7 or 8 or 9 or 10

12. Ultrasonography/
13. Endosonography/
14. Imaging, Three-Dimensional/
15. (TPUS or EVUS or DEA or EFD or ultraso* or endosonogra* or echogra* or sonogra* or echodefecogra*).ti,ab
16. 12 or 13 or 14 or 15
17. Rectocele/
18. Intussusception/
19. Rectal Prolapse/
20. Rectal Diseases/
21. Rectum/
22. (rectocele or enterocele or sigmoidocele or intussusception or (rectal n1 prolapse) or anismus or (perineal n1 descent) or (pelvic n1 floor n1 descent) or (posterior n1 compartment)).ti,ab
23. 17 or 18 or 19 or 20 or 21 or 22
24. ODS.ti,ab
25. (obstruct* n1 (defecat* or defaecat*)).ti,ab
26. ((defecat* or defaecat* or evacuat* or anorect*) n3 (disorder* or dysfunct* or difficult*)).ti,ab
27. 24 or 25 or 26
28. 11 or 16
29. 23 or 27
30. 5 and 28 and 29

Appendix 5. Search strategy for Science Citation Index/Conference Proceedings Citation Index

Science Citation Index 1900 to 18.12.2019 / Conference Proceedings Citation Index 1970 to 18.12.2019
#1 TS=(EP or defecogra* or defaecogra* or proctogra* or colpocystodefecogr* or colpocystorectogra* or videoproctogra* or cystocolpoproctogra* or fluorosco*)
#2 TOPIC: (((magnetic NEAR/1 resonance) or MRI or MR))
#3 TS=(TPUS or EVUS or DEA or EFD or ultraso* or endosonogra* or echogra* or sonogra* or echodefecogra*)
#4 TS=((rectocele or enterocele or sigmoidocele or intussusception or (rectal NEAR/1 prolapse) or anismus or (perineal NEAR/1 descent) or (pelvic NEAR/1 floor NEAR/1 descent) or (posterior NEAR/1 compartment)))
#5 TOPIC: ((ODS or (obstruct* NEAR/1 (defecat* OR defaecat*))))
#6 TOPIC: (((defecat* or defaecat* or evacuat* or anorect*) NEAR/3 (disorder* OR dysfunct* OR difficult*)))
#7 (#3 OR #2)
#8 (#6 OR #5 OR #4)
#9 (#8 AND #7 AND #1)

Appendix 6. Study eligibility screening proforma

Inclusion Criteria:

A Study design:

☐ Cross sectional test accuracy study

☐ Cohort (prospective or retrospective) test accuracy study

☐ Comparison of the accuracy of tests or testing strategies in two different populations (e.g. RCT)

☐ Any other study where estimation of test accuracy was not the primary objective

B Participants:

☐ Female patients with symptoms of ODS

☐ Female patients with symptoms of pelvic floor dysfunction

☐ Study selecting both female and male patients with symptoms of ODS or pelvic floor dysfunction (NB test accuracy data on women only need to be retrieved).

☐ Study selecting both asymptomatic and symptomatic women (NB test accuracy data of women with symptoms only need to be retrieved).

C Variable index test:

☐ Transperineal ultrasound

☐ Perineal ultrasound

☐ Introital ultrasound

☐ Translabial ultrasound

☐ Endovaginal ultrasound

☐ Anorectal ultrasound

☐ Echodefaecography

☐ Dynamic anal endosonography

☐ Dynamic MRI

☐ Open-magnet Dynamic MRI

D Fixed index test:

☐ Conventional EP or its equivalents CCD, CCP, CCRG, ECCP

E Target condition:

☐ Rectocele

☐ Enterocele

☐ Intussusception

☐ Anismus

☐ Pelvic floor descent

Tick in the box of section A and B and C and D and E: Inclusion

Tick in the box in four out of five sections: Discussion

Exclusion Criteria:

A Study design:

☐ Case control study comparing patients with and without a target condition (rectocele, enterocele, intussusception, anismus and pelvic floor descent)

☐ Case reports

☐ Reviews

B Participants:

☐ Age < 18

☐ Only men

☐ Only asymptomatic patients

Tick in one of above boxed? -> Exclusion

Appendix 7. Standardised data extraction form

A STUDY CHARACTERISTICS:

1. STUDY IDENTIFICATION AND STUDY TYPE

DETAILS

(Continued)

Title	
Authors	
Year of Publication	
Journal	
Country in which study is conducted	
Period of data collection	
Objective	
Study design (select one)	☐ Cross sectional test accuracy study ☐ Cohort test accuracy study ☐ Comparison of the accuracy of tests or testing strategies in two different populations (e.g. RTC)

2. PATIENT SELECTION

A. DETAILS	
Describe methods of patient selection (cut and paste from paper if possible)	
Describe characteristics included patients; previous testing, presentation intended use of index test, and setting (cut and paste from paper if possible)	
If studies evaluate more than one index test, how were test allocated to individuals, or did each individual receive all index tests?	
Number of participants	Total included: Nr of eligible patients: Nr of excluded patients:
Referral route	
Setting	Secondary / Tertiary Single centre/ Multi centre
Eligibility criteria	
Exclusion criteria	
Participant recruitment	Prospective/ Retrospective
Age	Age mean:

(Continued)

	Age range:
Gender	Female n (%): Male n (%): Male/female ratio:
Ethnicity	
Co-morbidities	
Symptoms	ODS / Constipation / Prolapse / other symtoms of pelvic floor dysfunction, being: Percentage:
B. ASSESSMENT RISK OF BIAS	
Was a consecutive or random sample of patients enrolled?	Yes / No / Unclear
Was a case–control design avoided?	Yes / No / Unclear
Did the study avoid inappropriate exclusions?	Yes / No / Unclear
Could the selection of patients have introduced bias?	Concern: High / Low / Unclear *Low risk on bias: All signalling questions are answered with 'yes'* *Unclear risk on bias: One or more signalling questions are answered as 'unclear'* *and none with 'high'* *High risk on bias: Any of signalling questions is answered with 'no'.*
C. CONCERNS ABOUT APPLICABILITY	
Are there concerns that the included patients do not match the review question?	Concern: High / Low / Unclear Motivation:

3. INDEX TEST (MRI OR ULTRASOUND)

A. DETAILS MRI OR ULTRASOUND	
Describe the index test and how it was conducted and interpreted (cut and paste from paper if possible)	
Method of MRI or Ultrasound:	Name: Type of MRI/US-scanner (manufacturer): Image acquisition:

(Continued)

Use of contrast	rectal / vaginal / rectal and vaginal / none If yes, type of contrast and volume:
Position of patient	supine/ left-lateral / upright
Evacuation phase	Yes/ No
Operator characteristics (e.g. training)	
Imaging analysis	One observer/ two observers Discrepancy meeting: Yes/No Blinded: Yes/No/Unknown
Thresholds used to define positive and negative tests for each target condition (delete condition if not assessed)	Rectocele Definition: Cut-off value test positive: Enterocele Definition: Cut-off value test positive: Intussusception Definition: Cut-off value test positive: Anismus Definition: Cut-off value test positive: Pelvic floor descent Definition: Cut-off value test positive:
B. ASSESSING RISK OF BIAS	
Were the index test results interpreted without knowledge of the results of the other index test(s)?	Yes / No / Unclear
If a threshold was used, was it prespecified?	Yes / No / Unclear
Could the conduct or interpretation of the index test have introduced bias?	Concern: High / Low / Unclear *Low risk on bias: All signalling questions are answered with 'yes'* *Unclear risk on bias: One or more signalling questions are answered as 'unclear'* *and none with 'high'* *High risk on bias: Any of signalling questions is answered with 'no'.*

(Continued)

C. CONCERNS ABOUT APPLICABILITY	
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	Concern: High / Low / Unclear Motivation:

4. INDEX TEST (EVACUATION PROCTOGRAM)

A. DETAIL	
Describe evacuation proctography and how it was conducted and interpreted (cut and paste from paper if possible) *NB in RevMan these results are entered in the 'reference standard' domain of the ROB assessment, as this domain could not be removed. Note that EP was not taken as reference standard in the meta-analysis but as index test.*	
Method of EP	Specific method (name): Type of X-ray machine (manufacturer): Image acquisition:
Use of contrast (more options possible)	Rectal / vaginal / rectal and vaginal / none If yes, type of contrast and volume:
Position of patient	supine/ left-lateral / upright
Evacuation phase	Yes/ No
Operator characteristics (e.g. training)	
Imaging analysis	One observer/ two observers Discrepancy meeting: Yes/No Blinded: Yes/No/Unknown
Thresholds used to define positive and negative tests for each target condition (delete condition if not assessed)	Rectocele Definition: Cut-off value test positive: Enterocele Definition: Cut-off value test positive: Intussusception Definition: Cut-off value test positive:

(Continued)

	Pelvic floor descent Definition: Cut-off value test positive:
B. ASSESSING RISK OF BIAS	
Were the results of EP interpreted without knowledge of the results of the other index test(s)?	Yes / No / Unclear
If a threshold was used, was it prespecified?	Yes / No / Unclear
Could the reference standard, its conduct, or its interpretation have introduced bias?	Concern: High / Low / Unclear *Low risk on bias: All signalling questions are answered with 'yes'* *Unclear risk on bias: One or more signalling questions are answered as 'unclear'* *and none with 'high'* *High risk on bias: Any of signalling questions is answered with 'no'.*
C. CONCERNS ABOUT APPLICABILITY	
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Concern: High / Low / Unclear Motivation:

5. FLOW AND TIMING

A. DETAIL	
Describe any patients who did not receive the index tests or reference standard or who were excluded from the 2 x 2 table (refer to flow diagram)	
Describe the interval and any interventions between index tests and the reference standard	
B. ASSESSING RISK OF BIAS	
Was there an appropriate interval between index tests and reference standard?	Yes / No / Unclear
Did all patients receive a reference standard?	Yes / No / Unclear
Did all patients receive the same reference standard?	Yes / No / Unclear
Were all patients included in the analysis?	Yes / No / Unclear
Could the patient flow have introduced bias?	Concern: High / Low / Unclear

(Continued)

	Low risk on bias: All signalling questions are answered with 'yes' *Unclear risk on bias: One or more signalling questions are answered as 'unclear'* *and none with 'high'* *High risk on bias: Any of signalling questions is answered with 'no'.*

TEST ACCURACY DATA

1. Data extraction for entering in RevMan:

	Evacuation proctogram positive	**Evacuation proctogram negative**	
Index test positive	True positive: N=	False positive: N=	Total index test +ve: N=
Index test negative	False negative: N=	True negative: N=	Total index test -ve: N=
	Total disease +ve: N=	Total disease -ve: N=	Total number tested: N=

2. Data extraction for entering in Meta-analysis:

Use when two tests are performed:

Patern	EP	Test 2	e.g.	Number of patients
1	Positive	Positive	TP	
2	Positive	Negative	FN	
3	Negative	Positive	FP	
4	Negative	Negative	TN	
Total number of patients				

Use when more tests are performed:

Patern	EP	Test 2	Test 3	Test ...	Number of patients
1	Positive	Positive	Positive	...	
2	Positive	Positive	Negative	...	
3	Positive	Negative	Negative	...	
4	Positive	Negative	Positive	...	
5	Negative	Positive	Positive	...	
6	Negative	Negative	Positive	...	
7	Negative	Positive	Negative	...	
8	Negative	Negative	Negative	...	
....	...	...	...	...	

Please add tables for each target condition under evaluation

Please add tables if more than one cut-off value is used, or when different methods of performing index test were used (e.g. with or without rectal contrast)

Appendix 8. Assessment of methodological quality QUADAS-2

DOMAIN 1: PATIENT SELECTION

A. DESCRIPTION	
Review question: Women with Obstructed Defaecation Syndrome	
Describe methods of patient selection	Patient selection: Study design: Study objective: Inclusion criteria: Exclusion criteria:
Describe included patients (previous testing, presentation, intended use of index test, and setting)	Nr of included patients: Gender: Age: Symptoms: Ethnicity: Co-morbidities: Setting:

(Continued)

	Time period: Country study is conducted:
B. SIGNALLING QUESTIONS	
Was a consecutive or random sample of patients enrolled? *A study ideally should enrol a consecutive or random sample of eligible patients with suspected disease to prevent the potential for bias.*	Yes= If it is clearly stated a consecutive or random sample of eligible patients was enrolled in the study. No= If it is clearly stated a selected (non-consecutive or non-random) sample of patients was enrolled in the study (e.g. women with presence of a target condition on clinical examination) or patients were selected by convenience. Unclear = If the method of patient recruitment or sampling is not reported or we could not tell.
Was a case–control design avoided? *Studies enrolling participants with known disease and a control group without the condition may exaggerate diagnostic accuracy.*	Yes= If the study avoided implementation of two separate selection processes to sample patients with the target condition and patients without the target condition. No= If the study did not avoid implementation of two separate selection processes to sample patients with the target condition and patients without the target condition. Unclear= If the method of selection processes is not reported or is unclear. *We did not include any case-control studies because this design might lead to overestimation of accuracy, hence this question is answered yes for all studies.*
Did the study avoid inappropriate exclusions? *Studies that make inappropriate exclusions (for example, not including 'difficult-to-diagnose' patients) may result in overestimation of diagnostic accuracy.*	Yes= If inclusion/exclusion criteria were presented and all patients with ODS or suspected with target conditions were included. a. Exclusion criteria are formulated and they are appropriate (e.g. < 18 age, certain co-morbidities not affecting target condition, contra-indications for one of the tests, non-willingness to participate, previous prolapse surgery, inability to strain). b. No exclusion criteria are formulated, but recruitment is consecutive (meaning no patients are excluded). No= If exclusion criteria are formulated and are inappropriate (e.g. exclusion of patients with or suspected to have one of the target conditions, exclusion of patients who could have undergone imaging for ODS, exclusion of patients based on age, education level, ethnicity or other psychosocial factors) Unclear = If the study did not provide clear definition of the selection (inclusion or exclusion) criteria and 'no' judgement is not applicable. a. No exclusion criteria are formulated and patients sampling is not consecutive (patients were excluded but we do not know if it is inappropriate) b. No exclusion criteria are formulated and it is unknown if recruitment was consecutive (unknown if patients were excluded and if it was inappropriate) NB In case of a retrospective study: if it has formulated exclusion criteria of not having had the index test or reference standard, this has to be considered as domain 1 rather than excluded from analysis (domain 4) as this causes selection bias
C. RISK OF BIAS	
Could the selection of patients have introduced bias?	Low risk on bias: All signalling questions are answered with 'yes' Unclear risk on bias:One or more signalling questions are answered as 'unclear' and none with 'high' High risk on bias:Any of signalling questions is answered with 'no'.

(Continued)

D. CONCERNS ABOUT APPLICABILITY	
Test accuracy data available for female patients only?	Yes / No / Unclear
Test accuracy data available for patients with ODS symptoms only?	Yes / No / Unclear
Are there concerns that the included patients do not match the review question?	High concern: If the study population differed from the population defined in the review question in terms of demographic features and co-morbidity (e.g. male included, asymptomatic patients) Low concern: If the study includes only clinically relevant population that would have undergone index test in real practice and includes representative form of target condition (e.g. women with symptoms of ODS) Unclear concern:If this information was unclear

DOMAIN 2: INDEX TEST

A. DETAILS	
Review question	Any type of imaging that could identify the target conditions: rectocele, enterocele, intussusception, anismus and pelvic floor descent. Because of the design of this meta-analysis this section will always contain the results of EP and any other type of imaging (MRI and Ultrasound).
Describe the index test and how it was conducted and interpreted	Name index test: *Name* Details of conducting index test: *Manufacturer, type probe, Patient position, Use of contrast, Level of expertise* Imaging acquisition: *Evacuation phase* Imaging analysis: *Examiners (number, level of expertise, blinding)* Threshold test positivity: *For each target condition*

B. ASSESSING RISK OF BIAS	
Were the index test results interpreted without knowledge of the results of the other index test(s)? *Knowledge of the one index test results may influence interpretation of the other index test results. The potential for bias is related to the subjectivity of interpreting index test and the order of testing.*	Yes = If the operators performing or interpreting the index test were unaware of the results of reference standard or the index test was always performed and interpreted before results of the reference standard were known. No= If the operators performing or interpreting the index test were aware of the results of reference standard (e.g. results of the index test are ascertained retrospectively from patient notes once the reference standard result is known) Unclear= If it is not reported whether the index test was conducted without knowledge of the results of the index test, or whether the index test was completed before the reference standard was known.
Was the threshold for test positivity pre-specified?	Yes= If the threshold (known or unknown) was defined before execution or interpretation of the index test (e.g. if in a study authors explicitly state they used a threshold specified prior to testing with the index test, even if they don't explicitly state what the threshold is)

(Continued)

	No= If the threshold for a positive result was not defined prior to test execution (e.g. if the threshold was chosen based on index test results performed at various thresholds to find the threshold with the best sensitivity/specificity) Unclear = If it was unclear whether the used threshold was pre-specified or not
C. RISK OF BIAS	
Could the conduct or interpretation of the index test have introduced bias?	Low risk on bias:All signalling questions are answered with 'yes' Unclear risk on bias: One or more signalling questions are answered as 'unclear' and none with 'high' High risk on bias: Any of signalling questions is answered with 'no'.
D. CONCERNS ABOUT APPLICABILITY	
Variations in test technology, execution, or interpretation may affect estimates of its diagnostic accuracy.	
If a reference line was used, was it the PCL?	Yes / No / Unclear / Not applicable
For MRI was the scanner a Tesla 1.0 or higher?	Yes / No / Unclear / Not applicable
Are there concerns that the index test, its conduct, or its interpretation differ from the review question?	High/Low/Unclear

DOMAIN 3: REFERENCE STANDARD

A. DETAIL	
Review Question	No reference standard is available, hence it was not possible to answer these sections. Consequently these sections have been removed from the QUADAS-2 assessment tool. Because of the design of this meta-analysis the results of EP are assessed in domain 2: index test.
Describe the reference standard and how it was conducted and interpreted (cut and paste from paper if possible)	Not applicable
B. ASSESSING RISK OF BIAS	
Were the reference standard results interpreted without knowledge of the results of the index test? *Knowledge of the index test results may influence interpretation of the reference standard results. Potential for bias is related to the potential influence of previous knowledge on the interpretation of the reference standard.*	Not applicable

(Continued)

Is the reference standard likely to correctly classify the target condition? *Estimates of test accuracy are based on the assumptions that the reference standard is 100% sensitive and that specific disagreements between the reference standard and index test result from incorrect classification by the index test.*	Not applicable
Was the threshold for test positivity pre-specified?	Not applicable
C. RISK OF BIAS	
Could the reference standard, its conduct, or its interpretation have introduced bias?	Not applicable
D. CONCERNS ABOUT APPLICABILITY	
Are there concerns that the target condition as defined by the reference standard does not match the review question?	Not applicable

4. FLOW AND TIMING

A. DETAIL	
Review question	Less than 3 months between imaging techniques as target conditions can progress (arbitrary cut-off).
Describe any patients who did not receive the index tests or who were excluded from the analysis (refer to flow diagram)	Enrolment and exclusions (+ reasons): Nr analysed:
Describe the interval and any interventions between the index tests	Time interval (+ interventions) between index tests:
B. ASSESSING RISK OF BIAS	
Was there an appropriate interval between index tests?	Yes = If time interval was reported and was less than 3 months No = If time interval was reported and was more than 3 months Unclear = If time interval was not stated clearly, but authors' description allowed to assume that the interval was reasonably short.
Did all patients underwent EP? Did all patients receive the index test irrespective of the other index test results? *Verification bias occurs when only a proportion of the study group receives confirmation of the diagnosis by the reference standard, or if some patients receive a different reference standard. If the results of the index test influence the decision on whether to perform the reference standard or which reference standard is used, estimated diagnostic accuracy may be biased.*	Yes = If all participants underwent evacuation proctogram (however not necessarily as a reference standard) No= If not all participants underwent evacuation proctogram or if only a subset of participants had evacuation proctogram, but the information on this population was not available in isolation. Unclear= If this information was unclear.

(Continued)

Were all patients included in the analysis? Were withdrawals from the study explained? Were uninterpretable/intermediate test results reported? *All participants recruited into the study should be included in the analysis. A potential for bias exists if the number of patients enrolled differs from the number of patients included in the 2 x 2 table of results.*	Yes = If all the women were included in the analysis or if not all women were included in the analysis but: - the withdrawals did not meet inclusion criteria prior to execution of index test (contra-indications for index test, non-willingness to participate) - the withdrawals are explained, appropriate and at random (patient did not attend appointment, lost or incomplete data sets) - excluded results are reported as uninterpretable results (not able to strain/ evacuate, poor quality of image) No = If any patients were excluded from the analysis for inappropriate reasons or exclusions were not explained. Unclear = If this information was unclear
C. RISK OF BIAS	
Could the patient flow have introduced bias?	*Low risk on bias: All signalling questions are answered with 'yes'* *Unclear risk on bias: One or more signalling questions are answered as 'unclear' and none with 'high'* *High risk on bias: Any of signalling questions is answered with 'no'.*

Appendix 9. Statistical Analysis: Bayesian hierarchical latent class analysis

The meta-analysis was performed using a Bayesian approach to hierarchical Latent Class Analysis, as described in the article on comparative Bayesian meta-analysis of diagnostic studies by Menten and Lesaffre (Menten 2015). Specifically we applied the hierarchical Latent Class Model (model 4), that is suitable for an imperfect reference standard.

The following syntax presents the OpenBUGS model that we used, here for a setting with two tests for which the sensitivity and specificity are to be estimated, in addition with the PPV, NPV, DOR, LR+ and LR-. Probabilities to evaluate whether an index test was suitable as a replacement test for EP or as SpIN or SnOUT triage test, are based on the MCMC chain of comparisons of the sensitivities and specificities of the index test with EP (diffSe2 and diffSp2).

Models similar to the Bugs model presented below, but with the appropriate number of tests (e.g. six tests), were fit using OpenBUGS version 3.2.3, with 3 chains, each with 100,000 iterations (burning 50,000).

```
model{

for (i in 1:nPats){

status[i] ~ dbern(prev[study[i]]) # true status of person i
for(k in 1:nTests){

Y[i,k] ~ dbern(P[i,k])

}

# Y[i,k]: observed result for person i test k

# P[i,k] is the probability of positive test result for test k person i

# alphak: vector of length 2 with logit (sens) and logit (1-spec) for test k.

logit(P[i,1]) <- status[i] * alpha1[study[i],1] + (1-status[i]) * alpha1[study[i],2]

logit(P[i,2]) <- status[i] * alpha2[study[i],1] + (1-status[i]) * alpha2[study[i],2]

}
```

```
# across studies, alphak is bivariate norm. distributed, mean muk, covar matrix inverse (Rk)
for(j in 1:nStudy){
alpha1[j,1:2] ~ dmnorm(mu1[],R1[,])
alpha2[j,1:2] ~ dmnorm(mu2[],R2[,])
prev[j] ~ dbeta(1,1)
logitprev[j] <- log(prev[j]/(1-prev[j]))
}
mean.prevlogit <- mean(logitprev[])
prevalence <- 1/(1+exp(-mean.prevlogit))
mu1[1] ~ dnorm(0,.37)I(0,)
mu2[1] ~ dnorm(0,.37)I(0,)
mu1[2] ~ dnorm(0,.37)I(,0)
mu2[2] ~ dnorm(0,.37)I(,0)
R1[1:2,1:2] <- inverse(RI1[1:2,1:2])
RI1[1,1] <- pow(sigma1[1],2)
RI1[1,2] <- cov1
RI1[2,1] <- cov1
RI1[2,2] <- pow(sigma1[2],2)
cov1 <- corr1*sigma1[1]*sigma1[2] # covariance logit sens with logit (1-spec)
corr1 ~ dunif(-1,1)
sigma1[1] ~ dnorm(0,1)I(.001,5) # between-study sigma
sigma1[2] ~ dnorm(0,1)I(.001,5)
R2[1:2,1:2] <- inverse(RI2[1:2,1:2])
RI2[1,1] <- pow(sigma2[1],2)
RI2[1,2] <- cov2
RI2[2,1] <- cov2
RI2[2,2] <- pow(sigma2[2],2)
cov2 <- corr2*sigma2[1]*sigma2[2]
corr2 ~ dunif(-1,1)
sigma2[1] ~ dnorm(0,1)I(.001,5) # dunif(.001,5)
sigma2[2] ~ dnorm(0,1)I(.001,5) # dunif(.001,5)
# calculations
SENS[1] <- 1/(1+exp(-mu1[1]))
SPEC[1] <- 1/(1+exp( mu1[2]))
SENS[2] <- 1/(1+exp(-mu2[1]))
```

```
SPEC[2] <- 1/(1+exp( mu2[2]))

PPV[1] <- (SENS[1]*prevalence) /

(SENS[1]*prevalence + (1-prevalence)*(1-SPEC[1]))

NPV[1] <- (SPEC[1] *(1-prevalence))/(SPEC[1]*(1-prevalence) +

+ (prevalence*(1-SENS[1])))

PPV[2] <- (SENS[2] *prevalence)/

(SENS[2]*prevalence + (1-prevalence)*(1-SPEC[2]))

NPV[2] <- (SPEC[2] * (1-prevalence))/(SPEC[2]*(1-prevalence) +

+ (prevalence*(1-SENS[2])))

DOR[1] <- 100*SENS[1]*SPEC[1]/ (100*(1-SPEC[1])*(1-SENS[1]))

DOR[2] <- 100*SENS[2]*SPEC[2]/ (100*(1-SPEC[2])*(1-SENS[2]))

LRpos[1] <- 100*SENS[1]/ (100*(1-SPEC[1]))

LRpos[2] <- 100*SENS[2]/ (100*(1-SPEC[2]))

LRneg [1] <- (1-SENS[1])/SPEC[1]

LRneg [2] <- (1-SENS[2])/SPEC[2]

# difference between sensitivities/ specificities, for calculation of probabilities

diffSe2[1] <- SENS[2]-SENS[1]

diffSp2[1] <- SPEC[2]-SPEC[1]

}
```

Appendix 10. Standardised form: Assessment level of evidence according GRADE

Consider which accuracy outcome link most directly to clinical outcome:

Is it more important that the index test rules out or rules in a target condition?

What are the harms of false positives and false negatives?

- Harm of a missed diagnosis:

- Harm of further testing or treatment:

Target condition – Imaging technique

Table summarising all data

Study ID	Risk of Bias				Concerns	Heterogeneity				N patients
	P	I	I	F	Applicability	Position	Rectal contrast	Evacuation phase	Cut-off value	

1. Risk of Bias

Summary Risk of Bias per domain (N studies (N patients))

	High	Unclear	Low
Patient selection			
MRI or Ultrasound			
EP			
Flow and timing			

Reasons high risk of bias domain 1 and 4:

Concerns about selection bias? Yes/No, if yes please explain

Reasons high risk of bias domain 2 and 3:

Concerns about verification bias/interpretation bias? Yes/No, if yes please explain

Summary Risk of Bias per study (N studies (N patients))

	Number of studies	Number of domains
Low risk of bias		All 4 domains
Unclear risk of bias		1 domain: , 2 domains: , 3 domains: , 4 domains:
High risk of bias		1 domain: , 2 domains: , 3 domains: , 4 domains:

DTA measures with and without high risk of bias studies

	N studies	Sensitivity (95% CrI)	Specificity (95% CrI)
All studies			
High risk of bias excluded			

When outcome measure differs 10% or more it should be downgraded.

Differences close to 10% is defined as borderline and in combination with another borderline judgement the total level of evidence should be downgraded with one level.

Judgment sensitivity: ROB not serious / serious / very serious / borderline

Judgment specificity: ROB not serious / serious / very serious / borderline

Explanation:

2. Directness

Summary Applicability (N studies (N patients))

	Yes		No	
	N studies	N patients	N studies	N patients
Women only	…	… (...%)	…	… (...%)
ODS only	…	… (...%)	…	… (...%)

DTA measures with only women with ODS

	N studies	Sensitivity (95% CrI)	Specificity (95% CrI)
All studies			
Only women with ODS			

When outcome measure differs 10% or more it should be downgraded.

Differences close to 10% is defined as borderline and in combination with another borderline judgement the total level of evidence should be downgraded with one level.

Judgment sensitivity: Direct / Indirect / Borderline

Judgment specificity: Direct / Indirect / Borderline

Explanation:

3. Consistency

Forest plots summarising data:

Co-variates assessed: patient position, evacuation phase, rectal contrast, cut-off value

If heterogeneity could be explained by co-variates judgment does not need to be downgraded.

Unexplained heterogeneity should be downgraded 1 level. Borderline judgements should in combination with another borderline judgement downgrade the total level of evidence with one level.

If studies were all performed by the same research group this should be downgraded an extra level.

Judgment sensitivity: Consistent / Inconsistent / Borderline

Judgment specificity: Consistent / Inconsistent / Borderline

Explanation:

4. Precision

Sample size: … studies, … patients

Confidence intervals: sensitivity (… -….), specificity (… - ….)

What happens with FN and NPV if sensitivity was 10% overestimated by our analysis?

- Sensitivity … (95%CI …): NPV …: FN …. (in cohort of 1000) ….%

- Sensitivity … (95%CI …): NPV …: FN …. (in cohort of 1000) ….%

- From …. to …. post test probability of having the condition after a negative test.

What happens with FP and PPV if specificity was 10% overestimated by our analysis?

- Specificity …. (95%CI …): PPV ….: FP …. (in a cohort of 1000) ….%

- Specificity …. (95%CI …): PPV ….: FP …. (in a cohort of 1000) ….%

- From ….% to …% post test probability of not having the condition after a positive test.

If the impact of imprecision on clinical outcomes is negligible or if the demonstrated precision is sufficient to make the decision, the evidence should not be downgraded.

NB as a combination of symptom severity, clinical examination and test result determines if patient would have treatment or not; in most cases the test result not too much influence on patient's outcome.

Downgrading if:

- Sample size 3 studies or fewer, or

- Sample size 400 patients or fewer, or

- Confidence interval reach 0.50 or width > 0.30 (if not yet downgraded for inconsistency)

Judgment sensitivity: Precise / Imprecise

Judgment specificity:Precise / Imprecise

Explanation:

5. Overall quality of the evidence

	Sensitivity	Specificity
Risk of bias judgment		
Directness judgment		
Consistency judgment		
Precision judgment		
Overall judgment	High / Moderate / Low / Very Low	High / Moderate / Low / Very Low

Explanation:

High: We are very confident that the true effect lies close to that of the estimate of the effect.

Moderate: We are moderately confident in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different.

Low: Our confidence in the effect estimate is limited: the true effect may be substantially different from the estimate of the effect.

Very low: We have very little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of effect.

NB Publication bias; not included as not performed

NB Dose response association; not included as not valid for DTA meta-analysis

NB Existence of plausible unmeasured confounders; not assessed

NB Strength of association (i.e. magnitude of effect); imperfect reference standard taken in account in all analysis

HISTORY

Protocol first published: Issue 1, 2015
Review first published: Issue 9, 2021

CONTRIBUTIONS OF AUTHORS

Draft the protocol	IMA van Gruting, Stankiewicz A, R Thakar, GA Santoro, J. IntHout, AH Sultan
Develop a search strategy	IMA van Gruting, CCCG Trials Search Coordinator
Search for studies (usually 2 people)	IMA van Gruting, Stankiewicz A
Obtain copies of studies	IMA van Gruting
Select which studies to include (2 + 1 arbiter)	IMA van Gruting, Stankiewicz A, R Thakar (arbiter)
Extract data from studies (2 people)	IMA van Gruting, Stankiewicz A, R Thakar (arbiter)
Enter data into RevMan	IMA van Gruting
Carry out the analysis	IMA van Gruting, J. IntHout
Interpret the analysis	IMA van Gruting, Stankiewicz A, R Thakar, GA Santoro, J IntHout, AH Sultan
Draft the final review	IMA van Gruting, Stankiewicz A, R Thakar, GA Santoro, J IntHout, AH Sultan
Update the review	IMA van Gruting, Stankiewicz A, R Thakar, GA Santoro, J IntHout, AH Sultan

DECLARATIONS OF INTEREST

The review authors have no conflicts of interest and no financial ties to disclose.

SOURCES OF SUPPORT

Internal sources

- Mayday Childbirth Charity Fund, UK

  IMA van Gruting is funded by the Mayday Childbirth Charity Fund

External sources

- Non, Other

  No external sources of funding

DIFFERENCES BETWEEN PROTOCOL AND REVIEW

Secondary objectives:

As a secondary objective in the protocol we aimed to assess test accuracy of each test at prespecified thresholds, but this was not possible due to insufficient data. The data that were available were used in the investigation of heterogeneity.

Criteria for considering studies for this review:

We included studies recruiting women with and without symptoms, to be able to retrieve extra test accuracy data on women with ODS.

Test positivity of TPUS for rectocele was defined as > 0 cm depth rather than > 10 mm depth, to have the same cut-off value as rectocele for all imaging techniques.

Selection of studies:

Eligible articles not in the English language were not translated, but we contacted the authors to supply the required information.

Assessment of methodological quality:

A domain was not only deemed at 'high risk on bias' when all signalling questions were answered with 'no', but when any of the signalling questions was answered with 'no'.

Statistical analysis and data synthesis:

Instead of the pairwise LCA approach as described by Chu 2009, we applied the Bayesian hierarchical LCA approach as described by Menten 2015.

We defined criteria for replacement test and triage test to establish the role in the diagnostic pathway of the index test under evaluation.

Investigation of heterogeneity:

We were unable to perform all planned analyses, due to insufficient data.

Sensitivity analysis:

In the protocol we said that we would perform an additional sensitivity analysis to determine the effect of excluding studies that were flagged as possibly less appropriate for inclusion (when disagreement between authors could not be resolved). We did not have any such studies, so this analysis was not performed.

Besides the planned sensitivity analysis excluding studies at high risk of bias, we also performed a sensitivity analysis excluding studies with concerns about applicability, an analysis excluding studies published before 2010, and an analysis without studies that were excluded in one or more of the previous analyses, to minimise all potential risk of bias.

Summary of findings:

We assessed the overall quality of the evidence using GRADE, to aid healthcare workers and decision-makers with the interpretation of the results. We did not refer to GRADE in the protocol.

INDEX TERMS

Medical Subject Headings (MeSH)

Bayes Theorem; Defecation; Defecography; *Pelvic Floor Disorders [complications] [diagnostic imaging]; Ultrasonography

MeSH check words

Female; Humans