Systematic review and meta-analysis: defecography should be a first-line diagnostic modality in patients with refractory constipation.

Ugo Grossi¹²; Gian Luca Di Tanna³; Henriette Heinrich⁴; Stuart A Taylor⁵; Charles H Knowles¹; S Mark Scott¹.

¹Centre for Trauma and Surgery, and GI Physiology Unit, Barts and the London School of Medicine and Dentistry, Queen Mary University of London, London, UK.
²Fondazione Policlinico Universitario A. Gemelli IRCCS, Rome, Italy.
³Riskcenter - IREA, Department of Econometrics, Statistics and Applied Economics, Universitat de Barcelona, Spain.
⁴Department of Gastroenterology and Hepatology, University Hospital Zurich, Zurich, Switzerland.
⁵Centre for Medical Imaging, University College London, London, UK.

Corresponding author:

Mr Ugo Grossi
Centre for Trauma and Surgery
Barts and The London School of Medicine and Dentistry
Queen Mary, University of London
4 Newark Street
London, E1 2AT
United Kingdom
Tel: +44 20 7882 2638
Email: u.grossi@qmul.ac.uk

Word count: 5326
Abstract

Background: Defecography is considered the reference standard for assessment of pelvic floor anatomy and function in patients with a refractory evacuation disorder. However, the overlap of radiologically significant findings seen in chronic constipation (CC) and healthy volunteers (HV) is poorly defined.

Aim: To systematically review and meta-analyse rates of structural and functional abnormalities diagnosed by barium defecography and/or magnetic resonance imaging defecography (MRID) in patients with symptoms of CC and HV.

Methods: Electronic searches of major databases were performed without date restrictions.

Results: From a total of 1760 records identified, 175 full-text articles were assessed for eligibility. A total of 63 studies were included providing data on outcomes of 7,519 barium defecographies and 668 MRIDs in patients with CC, and 225 barium defecographies and 50 MRIDs in HV. Pathological high-grade (Oxford III and IV) intussuscepta and large (>4 cm) rectoceles were diagnosed in 23.7%(95%CI,16.8-31.4) and 15.9%(10.4-22.2) of patients, respectively. Enterocele and perineal descent were observed in 16.8% (12.7-21.4) and 44.4% (36.2-52.7) of patients, respectively. Barium defecography detected more intussuscepta than MRID (OR,1.52 [1.12-2.14];p=0.009). Normative data for both barium defecography and MRID structural and functional parameters were limited, particularly for MRID (only one eligible study).

Conclusions: Pathological structural abnormalities, as well as functional abnormalities, are common in CC patients. Since structural abnormalities cannot be evaluated using non-imaging test modalities (balloon expulsion and anorectal manometry) defecography should be considered first-line diagnostic test, if resources allow.
Keywords: barium defecography; evacuation proctography; magnetic resonance imaging; defecography; chronic constipation; evacuation disorder; structural abnormalities; functional abnormalities.
INTRODUCTION

Chronic constipation (CC) affects up to 14% of the general population in Western countries,\(^1\) with pathophysiology commonly accepted as an overlap between slow colonic transit and/or an evacuation disorder (ED).\(^2\) ED may result from ‘structural’ causes (e.g. intussusception, rectocele, enterocele) and/or ‘functional’ disorders (e.g. impaired recto-anal coordination) of the anorectal region.\(^3\) As symptoms alone do not reliably discriminate between CC subtypes, anorectal physiological testing and imaging are usually warranted in those patients with refractory symptoms.\(^4\) The balloon expulsion test (BET), anorectal manometry, and defecography represent the 3 main diagnostic modalities.\(^3\) BET and manometry are currently considered the first-line tests,\(^5\) but \textit{de facto} these do not provide any information on structural abnormalities that may impede evacuation.

Defecography is a radiologic technique still considered as the reference standard for the assessment of pelvic floor anatomy and function,\(^6\) given its capability to dynamically evaluate the rectum (and other pelvic organs) during simulated defecation.\(^7\) Its particular advantage over BET and manometry is that it enables characterisation of structural abnormalities.\(^8\)\(^9\) However, in the assessment of functional parameters (i.e. of recto-anal coordination during straining), there is still debate over which test should be considered the gold standard, especially in selecting patients who may be more likely responsive to biofeedback therapy. Indeed, current evidence has shown considerable disagreement between the results of the 3 diagnostic modalities.\(^9\)\(^10\)

\textbf{Historical perspective and terminology}

X-ray barium defecography was originally described in the 1950s, where spot films were taken in patients with CC.\(^11\)\(^12\) Methodologic improvements refined the technique,\(^13\)\(^-\)\(^18\) whereby it has become more routinely available since the 1980s.\(^19\)\(^\ 20\) Barium
defecography has also been conducted using simultaneous administration of contrast agents into other organs (e.g. bladder, vagina, small intestine, or peritoneum) to overcome its inability to depict the perirectal soft tissues. However, these steps inevitably increase the invasiveness of the examination, which also involves exposing the patient to ionizing radiation. Consequently, interest in utilising magnetic resonance imaging as an alternative modality by which to perform defecography (MRID) has been increasing since its first report in 1991.

Regardless of technique, a consistent criticism of defecography is the acknowledged overlap between health and disease, hampered by a paucity of normative data, which challenges our ability to define ‘true’ abnormalities. Even terminology is far from being universally accepted, given the numerous technical variations proposed and the plethora of synonyms of defecography since its conception: ‘cineradiographic defecography’, ‘cinedefecography’, ‘evacuating’ or ‘evacuation’ ‘proctography’, ‘defecation’ or ‘defecating’ ‘proctography’, ‘videodefecography’, and ‘videoproctography’. The term ‘defecography’ has been most commonly reported (~60% of all published articles), and was initially proposed by Mahieu to more clearly imply that the physiological act of ‘defecation’ is examined in dynamic conditions analogous to the investigation of deglutition or micturition. For the sake of simplicity, we have adopted its use in the manuscript to designate both X-ray barium and magnetic resonance imaging techniques.

**Technique**

**X-ray barium defecography**

The first symposium on barium defecography, in 1988, brought together the knowledge of 10 experts from 6 tertiary centres across the world (United Kingdom, Sweden, The Netherlands, Belgium, United States, and Canada). Considerable variation in technique
was immediately apparent, in terms of patient position, bowel preparation, consistency of
contrast materials, types of radiolucent commode, definitions of normality and abnormality
and their clinical implications. Surprisingly, most of these variations have continued to the
present, as discussed below.

Beside the rectum, opacification can be extended to vagina, bladder and/or small bowel.
A viscous contrast material is routinely used to achieve a consistency similar to stool.
Proprietary commercial formulations are available (e.g. 100% weight for volume barium
sulfate agent Anotrast® [E-Z-EM, Westbury, NY]) to be administered via a caulking gun.
Alternatively, homemade physiological pastes are preferred by some institutions using
organic ingredients (e.g. potato and oatmeal mixes), which then have barium added. Thick
barium (10 ml to 30 ml) or soaked tampons have been utilized for vaginal opacification.
Barium suspension (100 ml to 500 ml) has also been administered orally for small bowel
opacification. In 4 studies, bladder opacification was achieved using 50-250 ml of
iodinated, radiopaque contrast medium.\textsuperscript{34-37} Conventional barium defecography exposes
patients to a mean radiation equivalent dose of 0.5-5.0 millisieverts and a gonadal
(equivalent) radiation dose of approximately 20-25 millisieverts in female patients.\textsuperscript{8,38,39}

\textbf{Magnetic resonance imaging defecography (MRID)}

The role of MRID in the evaluation of pelvic floor disorders has been less extensively
investigated. The obvious advantage over barium defecography is the ability of MRID to
simultaneously assess the three pelvic compartments with good accuracy, without ionising
radiation, and with limited invasiveness or discomfort.\textsuperscript{40} However, MRID is usually
performed using closed 1.5 Tesla magnets with the patient supine, which is often criticised
as non-physiological.\textsuperscript{41,42} Although MRID can also be performed with an open magnet
(thus allowing a physiological sitting position), initial comparative studies between open
and closed systems showed reasonable concordance of findings, hence validating the
widespread use of the latter. Nevertheless, more recent evidence suggests that MRID using closed-magnet systems, with the patient in a supine position, overestimates the grade of the dynamic descent of the pelvic floor and, at the same time, may result in underestimation of the severity of all disorders compared to open-magnet with the patient in a sitting position.

As with barium defecography, MRID lacks technical standardization of equipment, available sequences and rectal contrast agents. Examination without rectal filling has gained increasing popularity, since severe dysfunctions can be disclosed at maximal straining without the need of an evacuation phase. As such, attempts to strain can be repeated several times to optimize capture of structural abnormalities (e.g. rectocele). Only recently, a panel of experts from the European Society of Urogenital Radiology (ESUR) and the European Society of Gastrointestinal and Abdominal Radiology (ESGAR) convened to define the minimum prerequisites to obtain a state-of-the-art MR examination of the pelvic floor. One of the key point was that static, dynamic and evacuation sequences should be generally performed. However, since all panelists were using MR with a conventional closed-magnet, procedural and technical aspects of pelvic floor imaging was focused to this type of system.

Objectives

Primary objectives

1. In patients and healthy volunteers (HV), to determine the rates (diagnostic yield) of structural abnormalities diagnosed by defecography, with a focus on intussusception and rectocele.

2. In patients and HV, to determine the rates (diagnostic yield) of functional abnormalities diagnosed by defecography, with a focus on dyssynergic defecation.
Secondary objectives

1. In patients, to determine whether differences exist in rates of main diagnosis between X-ray barium defecography and magnetic resonance imaging defecography (MRID).

2. In patients, to determine the rates (diagnostic yield) of structural abnormalities diagnosed by defecography when rigid normative data ranges are applied as cut offs.

METHODS

The authors developed the protocol for review, detailing pre-specified methods of analysis and eligibility of the studies in line with 2009 Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) guidance.48

Study characteristics

Search term definitions were inclusive, promoting a wide search of studies reporting diagnostic yield of barium defecography and/or MRID in patients with CC and/or HV. As the definition of CC is not standardized and uniformly applied,49 all common terms encompassing problematic defecation were used [Appendix 1]. Studies were eligible regardless of whether they were retrospective or prospective in design, controlled or uncontrolled. They were eligible if they provided extractable data on the prevalence of radiological abnormalities (structural and/or functional) on barium defecography and/or MRID. Studies were ineligible for inclusion if they described the use of defecography in patients suffering from bowel complaints other than constipation (e.g. fecal incontinence), and in whom constipation did not represent the primary presenting complaint. Similarly, studies were excluded if outcomes could not be segregated for the index population (i.e.
coexistent constipation and fecal incontinence or anal pain, or gynecologic complaints in women, where data were not stratified) and if the population affected by CC was enriched \textit{a priori} based on clinical and/or radiological confirmation of specific abnormalities (either structural or functional). Studies without clear definition of radiological abnormalities were also excluded (i.e. definitions were neither reported in the text nor referenced in the methods).

A minimum population sample of 40 adult subjects (index population) was imposed for eligibility. This pragmatic threshold was taken to exclude case reports and small case series that often reported on early experience with the techniques.

\textbf{Report characteristics}

Any publication date was eligible to the date of the final search performed on 05 November 2017. Due to the large number of studies retrieved, it was decided to include only studies with full-text in the English language. This approach is supported by the evidence that literature searches limited to English language publications do not affect the quality of systematic reviews.\textsuperscript{50} Only peer-reviewed publications reporting primary data were eligible. Thus reviews, editorials, and letters were excluded at the screening stage. Conference abstracts and proceedings were also excluded.

\textbf{Information sources and study selection}

The authors performed a comprehensive search of the literature using Medline (PubMed) and EMBASE and hand-searching using all common search terms encompassing problematic defecation\textsuperscript{51} and defecography with synonymous variants [Appendix 1]. Reference lists of all full-texts were hand-selected for any additional studies.
Data extraction

Screening was performed at the abstract level by the junior authors (UG and HH), excluding studies not meeting eligibility criteria where these could be readily determined from the abstract alone. Full-text copies of all remaining studies were also obtained and assessed by the junior authors, who were un-blinded to the names of studies, authors, institutions or publications (Figure 1). Disagreement regarding inclusion was resolved by a senior author (SMS). Study characteristics and outcome data were extracted independently by the junior authorship team onto a Microsoft Excel spreadsheet (XP professional edition; Microsoft Corp, Redmond, Washington, USA), with disagreements resolved by consensus.

The following data were extracted for each study: publication year, country of origin, reason for exclusion, total number of patients, number of females, number of patients with constipation as primary complaint, number of HV, number of controls, mean or median age, bowel preparation prior to start the procedure, volumes of rectal, oral, vaginal, and/or vesical contrast, definitions and prevalence of structural (i.e. internal [stratified as recto-rectal and recto-anal where applicable] and external rectal prolapse, rectocele [total and >4 cm], enterocele, megarectum, dynamic perineal descent and cut-off used for definition) and functional abnormalities (i.e. a) paradoxical or incomplete relaxation of the puborectalis muscle, b) incomplete opening of the anal canal, and/or c) incomplete rectal evacuation), and study quality using the component 1 (10-criterion checklist) proposed by Guo et al.\textsuperscript{52} to indicate the extent to which a case series presented traditional features of a statistical hypothesis-testing paradigm (Suppl. Figure 1).
Meta-analysis and subgroup analyses

The proportion of structural and functional abnormalities in each study (obtained by dividing the number of abnormality by the total number of CC and/or HV) was combined to give a pooled prevalence for all studies. For structural abnormalities, this was performed irrespective of the criteria used to define their presence. However, calculation of pooled prevalence was made according to specific diagnostic criteria for pathologically significant intussusception and rectocele (i.e. a magnitude not seen in studies in HV). Quantitative heterogeneity between studies was assessed using the $I^2$ statistic, acknowledging that the use of a specific threshold might lead to potentially misleading interpretation. Both fixed and random models results were presented, providing comments on the random effects when the $I^2$ was higher than 50% (commonly referred to as substantial heterogeneity). Subgroup analyses were conducted according to volume of rectal contrast administered (i.e. ≤150 ml, 151-200 ml, and >200 ml or defecatory desire volume) and diagnostic criteria for functional abnormalities (see Data extraction). The prevalence of enterocele was compared according to use of oral contrast using an odds ratio (OR) with a 95% confidence interval (CI). The pooled proportions (prevalence of structural and functional abnormalities) were calculated using the Freeman-Tukey double arcsine transformation method; odds ratio (OR) and 95% confidence interval (CI) was calculated for the following comparisons: oral vs non-oral contrast studies subgroups (for prevalence of enterocele), barium defecography vs MRID (for prevalence of functional abnormalities in non-controlled studies, for prevalence of structural abnormalities and intolerance in controlled studies), different diagnostic criteria within barium defecography and MRID (for prevalence of functional abnormalities). All data were pooled using fixed- and random-effects models with prevalence results reported along with 95%CI. All analyses were performed in STATA.
15 (StataCorp LLC, College Station, TX, USA), using the *Metaprop* function to obtain the pooled prevalence.\(^{54}\)

### SUMMARY OF SEARCH RESULTS AND STUDY QUALITY

#### Study selection

From a total of 1760 records identified, 1757 were screened after duplicates removed, 1582 of which were excluded. The database search yielded a total of 175 articles for full text review (Figure 1). Specific exclusions after full-text review included 45 studies where prevalence of radiological abnormalities in patients with CC could not be segregated from those suffering from other defecatory disorders (fecal incontinence, anal pain, gynecologic complaints in women); 24 studies where the population affected by CC was confirmed to be less than 40 patients; 13 studies where the population affected by CC was selected *a priori* based on clinical and/or radiological confirmation of specific abnormalities; 7 studies where CC did not represent the primary presenting complaint; 7 studies with no clear definition of radiological abnormalities; 7 publications reporting a patient cohort that overlapped with other studies; 5 reviews; 3 studies with no extractable data on the prevalence of radiological abnormalities; 1 study where peritoneography was performed prior to defecography.

Overall, 63 studies published between 1984 and 2017 contributed to the systematic review, providing data on the diagnostic yield of 7,519 barium defecographies (range, 40-896 per study) and 668 MRIDs (range, 40-188 per study) in patients with CC, and 225 barium defecographies (range, 8 to 47 per study) and 50 MRIDs (n = 1 study) in HV (Table 1). Overall, only 9 (14%) studies were controlled using either healthy (n = 2)\(^{55,56}\) or non-healthy (n = 3; e.g. patients presenting for investigation of other complaints) volunteers,\(^{57-59}\) or a combination of both subjects (n = 4).\(^{60-63}\)
Of the 63 articles included, 53 observational studies reported on barium defecography, 5 on MRID, and 5 on direct comparisons between the two techniques. A total of 45 studies originated from European centers, 9 from the USA and 9 from other countries.
A total of 4 studies were exclusively conducted in HV\textsuperscript{8,19,20,64} and 59 studies in patients with CC. A meta-analysis was restricted to the latter group; considering the methodological heterogeneity of the 4 HV studies it was deemed not appropriate to pool the results for these. Overall, 50 studies reported on the prevalence of structural and/or functional abnormalities using barium defecography, 4 using MRID, and 5 using both techniques (comparative studies).
A total of 60/63 (95%) studies reported a male/female ratio. Of these, only 1 study exclusively recruited male patients with CC.\textsuperscript{65} The other 59 studies reported outcomes on 6,334 (87%; median 72, interquartile range [IQR] 52-113) females and 931 (median 12, IQR 0-22) males among CC patients (n=55 studies), and 110 (55%; median 27, IQR 24-33) females and 89 (median 23, IQR 19-25) males among HV (n=4 studies).

**Table 1:** All studies included in the systematic review.

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>Technique</th>
<th>Patients* (N)</th>
<th>Controls (N)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mahieu\textsuperscript{19}</td>
<td>1984</td>
<td>Belgium</td>
<td>BD</td>
<td>0</td>
<td>56\textsuperscript{†}</td>
</tr>
<tr>
<td>Mahieu\textsuperscript{66}</td>
<td>1984</td>
<td>Belgium</td>
<td>BD</td>
<td>144</td>
<td>0</td>
</tr>
<tr>
<td>Bartolo\textsuperscript{63}</td>
<td>1988</td>
<td>UK</td>
<td>BD</td>
<td>49</td>
<td>25\textsuperscript{§}</td>
</tr>
<tr>
<td>Shorvon\textsuperscript{20}</td>
<td>1989</td>
<td>Canada</td>
<td>BD</td>
<td>0</td>
<td>47</td>
</tr>
<tr>
<td>Felt-Bersma\textsuperscript{67}</td>
<td>1990</td>
<td>The Netherlands</td>
<td>BD</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>Poon\textsuperscript{27}</td>
<td>1991</td>
<td>UK</td>
<td>BD</td>
<td>63</td>
<td>0</td>
</tr>
<tr>
<td>Nielsen\textsuperscript{68}</td>
<td>1993</td>
<td>Denmark</td>
<td>BD</td>
<td>93</td>
<td>0</td>
</tr>
<tr>
<td>Siproudhis\textsuperscript{69}</td>
<td>1993</td>
<td>France</td>
<td>BD</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>Ger\textsuperscript{70}</td>
<td>1993</td>
<td>USA</td>
<td>BD</td>
<td>116</td>
<td>0</td>
</tr>
<tr>
<td>Klauser\textsuperscript{71}</td>
<td>1994</td>
<td>Germany</td>
<td>BD</td>
<td>97</td>
<td>0</td>
</tr>
<tr>
<td>Lee\textsuperscript{72}</td>
<td>1994</td>
<td>Taiwan</td>
<td>BD</td>
<td>55</td>
<td>0</td>
</tr>
<tr>
<td>Karlbom\textsuperscript{73}</td>
<td>1995</td>
<td>Sweden</td>
<td>BD</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>Halligan\textsuperscript{74}</td>
<td>1995</td>
<td>UK</td>
<td>BD</td>
<td>74</td>
<td>0</td>
</tr>
<tr>
<td>Halligan\textsuperscript{75}</td>
<td>1996</td>
<td>UK</td>
<td>BD</td>
<td>60</td>
<td>0</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Country</td>
<td>Type</td>
<td>BD</td>
<td>MRID</td>
</tr>
<tr>
<td>------------------</td>
<td>------</td>
<td>-------------</td>
<td>------</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>Agachan</td>
<td>1996</td>
<td>USA</td>
<td>BD</td>
<td>232</td>
<td>0</td>
</tr>
<tr>
<td>Schouten</td>
<td>1997</td>
<td>The Netherlands</td>
<td>BD</td>
<td>170</td>
<td>29§</td>
</tr>
<tr>
<td>Wiersma</td>
<td>1997</td>
<td>The Netherlands</td>
<td>BD</td>
<td>248</td>
<td>14†</td>
</tr>
<tr>
<td>Pfeifer</td>
<td>1997</td>
<td>USA</td>
<td>BD</td>
<td>100</td>
<td>0</td>
</tr>
<tr>
<td>Tsiaoussis</td>
<td>1998</td>
<td>Greece</td>
<td>BD</td>
<td>162</td>
<td>44†</td>
</tr>
<tr>
<td>Glia</td>
<td>1998</td>
<td>Sweden</td>
<td>BD</td>
<td>134</td>
<td>0</td>
</tr>
<tr>
<td>Mellgren</td>
<td>1998</td>
<td>Sweden</td>
<td>BD</td>
<td>112</td>
<td>0</td>
</tr>
<tr>
<td>Karlborn</td>
<td>1999</td>
<td>Sweden</td>
<td>BD</td>
<td>215</td>
<td>30†</td>
</tr>
<tr>
<td>Spazzafumo</td>
<td>1999</td>
<td>Italy</td>
<td>BD</td>
<td>316</td>
<td>0</td>
</tr>
<tr>
<td>Barthet</td>
<td>2000</td>
<td>France</td>
<td>BD</td>
<td>154</td>
<td>25§</td>
</tr>
<tr>
<td>Faucheron</td>
<td>2000</td>
<td>France</td>
<td>BD</td>
<td>43</td>
<td>0</td>
</tr>
<tr>
<td>Goh</td>
<td>2000</td>
<td>UK</td>
<td>MRID</td>
<td>0</td>
<td>50</td>
</tr>
<tr>
<td>Dailianas</td>
<td>2000</td>
<td>Greece</td>
<td>BD</td>
<td>49</td>
<td>22</td>
</tr>
<tr>
<td>Stojkovic</td>
<td>2000</td>
<td>UK</td>
<td>BD</td>
<td>136</td>
<td>0</td>
</tr>
<tr>
<td>Mibu</td>
<td>2001</td>
<td>Japan</td>
<td>BD</td>
<td>46</td>
<td>0</td>
</tr>
<tr>
<td>Gosselink</td>
<td>2001</td>
<td>The Netherlands</td>
<td>BD</td>
<td>80</td>
<td>60§</td>
</tr>
<tr>
<td>Savoye-Collet</td>
<td>2003</td>
<td>France</td>
<td>BD</td>
<td>52</td>
<td>0</td>
</tr>
<tr>
<td>Yeh</td>
<td>2003</td>
<td>USA</td>
<td>BD</td>
<td>261</td>
<td>0</td>
</tr>
<tr>
<td>Karlborn</td>
<td>2004</td>
<td>Sweden</td>
<td>BD</td>
<td>127</td>
<td>0</td>
</tr>
<tr>
<td>Dvorkin</td>
<td>2005</td>
<td>UK</td>
<td>BD</td>
<td>896</td>
<td>0</td>
</tr>
<tr>
<td>Renzi</td>
<td>2006</td>
<td>Italy</td>
<td>BD</td>
<td>420</td>
<td>0</td>
</tr>
<tr>
<td>Soares</td>
<td>2009</td>
<td>Brazil</td>
<td>BD</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>Brusciano</td>
<td>2009</td>
<td>Italy</td>
<td>BD</td>
<td>84</td>
<td>10</td>
</tr>
<tr>
<td>Murad-Regadas</td>
<td>2009</td>
<td>USA</td>
<td>BD</td>
<td>255</td>
<td>0</td>
</tr>
<tr>
<td>Morandi</td>
<td>2010</td>
<td>Italy</td>
<td>BD</td>
<td>567</td>
<td>0</td>
</tr>
<tr>
<td>Baek</td>
<td>2010</td>
<td>South Korea</td>
<td>BD</td>
<td>136</td>
<td>0</td>
</tr>
<tr>
<td>Mohammed</td>
<td>2010</td>
<td>UK</td>
<td>BD</td>
<td>200</td>
<td>0</td>
</tr>
<tr>
<td>Vitton</td>
<td>2011</td>
<td>France</td>
<td>C</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>Martellucci</td>
<td>2011</td>
<td>Italy</td>
<td>BD</td>
<td>54</td>
<td>0</td>
</tr>
<tr>
<td>Regadas</td>
<td>2011</td>
<td>Multicentre</td>
<td>BD</td>
<td>86</td>
<td>0</td>
</tr>
<tr>
<td>Ribas</td>
<td>2011</td>
<td>Spain</td>
<td>BD</td>
<td>106</td>
<td>0</td>
</tr>
<tr>
<td>Bordeianou</td>
<td>2011</td>
<td>USA</td>
<td>BD</td>
<td>123</td>
<td>0</td>
</tr>
<tr>
<td>Viscardi</td>
<td>2012</td>
<td>Italy</td>
<td>BD</td>
<td>46</td>
<td>0</td>
</tr>
<tr>
<td>Pilkington</td>
<td>2012</td>
<td>UK</td>
<td>C</td>
<td>42</td>
<td>0</td>
</tr>
<tr>
<td>Alves-Ferreira</td>
<td>2012</td>
<td>USA</td>
<td>BD</td>
<td>58</td>
<td>0</td>
</tr>
<tr>
<td>Piloni</td>
<td>2013</td>
<td>Italy</td>
<td>MRID</td>
<td>105</td>
<td>0</td>
</tr>
<tr>
<td>Seong</td>
<td>2013</td>
<td>South Korea</td>
<td>BD</td>
<td>96</td>
<td>0</td>
</tr>
<tr>
<td>Adusumilli</td>
<td>2013</td>
<td>UK</td>
<td>BD</td>
<td>64</td>
<td>0</td>
</tr>
<tr>
<td>Kashyap</td>
<td>2013</td>
<td>USA</td>
<td>BD</td>
<td>45</td>
<td>0</td>
</tr>
<tr>
<td>Andrade</td>
<td>2014</td>
<td>Portugal</td>
<td>BD</td>
<td>300</td>
<td>0</td>
</tr>
<tr>
<td>Palit</td>
<td>2014</td>
<td>UK</td>
<td>BD</td>
<td>0</td>
<td>46</td>
</tr>
<tr>
<td>Li</td>
<td>2015</td>
<td>China</td>
<td>MRID</td>
<td>56</td>
<td>0</td>
</tr>
<tr>
<td>Heinrich</td>
<td>2015</td>
<td>Switzerland</td>
<td>MRID</td>
<td>188</td>
<td>0</td>
</tr>
<tr>
<td>Kassis</td>
<td>2015</td>
<td>USA</td>
<td>BD</td>
<td>61</td>
<td>0</td>
</tr>
</tbody>
</table>
| Patient suffering from chronic constipation (NB: the number may differ from the original total sample size). BD: barium defecography; MRID: magnetic resonance imaging defecography; C: studies comparing BD vs. MRID; † Non-healthy controls; § Combination of non-healthy and healthy controls.

| Study quality |
| The 63 included studies were all observational with no experimental allocation to tests. The majority of studies were retrospective in nature (56%). One further limitation was blinding, with only 28% of studies stating that all defecographic images were reviewed by assessors who were not aware of the patient history to minimize observer bias (Suppl. Figure 1). Duration of follow-up was reported in 17/28 (61%) prospective studies with a mean (standard deviation) of 23 (14) months of follow-up. Interestingly, prevalence of the 2 most common truly pathological structural abnormalities (i.e. recto-anal intussusception and large [>4 cm] rectocele – discussed in detail below) was higher in prospective than retrospective studies (33.7 [21.0-47.6] vs. 17.1 [10.6-24.7], and 23.1 [14.5-32.9] vs. 11.6 [5.6-19.2], respectively).

Substantial heterogeneity was seen in all pooled prevalence.

| Structural abnormalities |
| Patients |

*Intussusception*

Pooled prevalence of intussusception on barium defecography was 36.8% (95%CI, 31.7-42.0) in patients with CC, and affected up to one third of patients with a clinically confirmed diagnosis. |
rectocele (Suppl. Figure 2). Similar rates of intussusception were observed on MRID (34.5 [21.9-48.3; based on 9 studies]) (Suppl. Figure 3).

The definition of intussusception was reported and/or referenced in a total of 40/63 (63%) studies. Among these, 3 main grading systems were used to define the severity of intussusception, originally described as “unilateral or circumferential infolding of the rectum during straining”. A total of 7 studies adopted the classification proposed by Shorvon et al., which identifies 7 degrees of intussusception, with grades 1 to 4 inclusive being intra-rectal (1 and 2: <3 mm; unilateral or circumferential, respectively; 3 and 4: >3 mm; unilateral or circumferential, respectively), 5 and 6 intra-anal (the leading edge of the infolding impinges onto or into the anal canal, respectively), and 7 representing an external rectal prolapse.

Among the 40 studies reporting definitions of intussusception, a total of 24 recognized intussuscepta as either unilateral or circumferential, whereas 16 studies regarded only circumferential intussuscepta as a truly abnormal finding. Among the former group, only 6 studies utilized specific cut-offs to determine the significance of the infolding: any fold “more than a wrinkling of the mucosa” (n=1), ≥3 mm (n=1), >4 mm (n=2), or >1 cm (n=2). When only reported if circumferential, intussuscepta were broadly stratified into intra-rectal, intra-anal, and external rectal prolapse, as originally described by Karlbom et al. Only 1 study adopted the more recent Oxford Prolapse Grade system to recognise intussuscepta, with the leading edge of the infolding descending no lower than proximal limit of the rectocele (grade I), or into the level of the rectocele but not onto sphincter/anal canal (grade II), or onto sphincter/anal canal (grade III), or into sphincter/anal canal (grade IV), or protruding from the anus (grade V) (Figure 2).

Rectocele
Pooled prevalence of a rectocele on barium defecography was 54.1% (95%CI, 48.0-60.2) in patients with CC (Suppl. Figure 4).

The definition of rectocele was reported and/or referenced in a total of 41/63 (65%) studies. Rectocele has traditionally been defined as an outpouching of the rectal wall on defecation. A total of 17/41 studies defined a cut-off of rectocele depth to establish the diagnosis: 2 cm (n=9), 2.5 cm (n=1), 3 cm (n=4), or 4 cm (n=3). However, the approach adopted to calculate rectocele size during maximum straining has been detailed in only 9 of these as: a) the ‘maximum depth of the bulge beyond the expected and extrapolated line of the anterior rectal wall’ (n=2); the ‘distance between the maximal anterior outbulge and b) the axis of the anal canal’ (n=2); or c) ‘a line through aspect of anorectal junction’ (n=1) or ‘a line drawn parallel to the center of the anal canal during straining’ (n=2); d) ‘outpouching of the anterior rectal wall ahead of rectovaginal septum, persisting on incomplete evacuation’ (n=2).

The amount of contrast retained within the rectocele has been reported as a measure of clinical significance in only 9/41 studies.

**Enterocoele**

Among the 27/59 (46%) studies reporting its prevalence in CC, enterocele affected a larger proportion of patients in studies describing the use of oral contrast (n=11) compared to those that did not (n=16) (20.4% [95%CI, 15.6-25.6] vs. 14.4% [8.8-21.1%], respectively; p=0.155), with an overall prevalence of 16.8% (95% CI, 12.7-21.4) on barium defecography and 15.8% (95% CI, 7.6-26.1) on MRID (Figure 3).

Enterocele has traditionally been defined as a herniation of the posterior cul-de-sac downward between the vagina and rectum. The hernia may contain small bowel or sigmoid colon. In the latter case, it is more commonly defined ‘sigmoidocele’.
Shorvon description of enterocele as an ‘indentation of posterior vaginal wall and anterior rectal wall’, 20 various defecographic definitions have been provided, with the simplest including ‘external compression of the anterior rectal wall during straining’; 72 or ‘contrast filled loops between rectum and vagina in women, and anterior to the rectum in men’; 57 or ‘semilunar defect in rectum during straining’. 83 Studies using MRID provide more accurate definitions, such as ‘small bowel within the rectovaginal septum that reached or crossed the junction of the upper one third and distal two thirds of the vagina’; 81 or ‘herniation of the peritoneal sac into the rectogenital space below the pubo-coccygeal line’. 106

A definition of enterocele and/or sigmoidocele was provided in only 22/63 (35%) studies. Small bowel opacification is pivotal to making a definitive diagnosis of enterocele with standard barium defecography, otherwise it is difficult to determine whether a widened rectovaginal space is due to a herniated mesentery or a prolapsed uterus, rather than enterocele. 113 However, administration of oral contrast was reported in only 19/63 (30%) studies.

_Perineal descent_

Descending perineum syndrome was first defined by Parks et al. in 1966 114 as an excessive ballooning of the perineum below the bony outlet of the pelvis associated with symptoms of ED, rectal pain, mucus discharge and/or rectal bleeding. Although measurement of perineal descent has been extensively reported in studies, there is poor consensus on definitions and pathophysiological implications. Lack of standardization comes from which position of the perineum should be measured, whether at rest (static) or during straining (dynamic). Even for the latter, various cut-off values have been used, ranging from 2 to 6 cm, 67 88 making estimate and comparison of prevalence rates very difficult (Table 2). Also, anatomical/fixed reference points vary among studies and include:
pubococcygeal line; upper surface of the commode; ischial tuberosity; a water-filled ring.

Pooled prevalence of significant dynamic perineal descent in patients with CC was 44.4% (95% CI, 36.2-52.7) on barium defecography and 43.6 (26.6-61.3) on MRID (Table 3; Suppl. Figure 5).

**Table 2.** Definitions and prevalence of dynamic perineal descent on barium defecography.

<table>
<thead>
<tr>
<th>Cut off (cm)</th>
<th>No. studies*</th>
<th>No. defecographies</th>
<th>Reference point from ARJ on straining</th>
<th>Pooled prevalence (%), 95%CI</th>
<th>( \hat{I}^2 ) (%)</th>
<th>( p ) value</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>2</td>
<td>99</td>
<td>PCL(^{94}) Water-filled ring(^{67})</td>
<td>54.3 (44.3-64.1)</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>7</td>
<td>884</td>
<td>PCL(^{27,72,78,80,115}) ARJ at rest(^{84}) IT(^{79})</td>
<td>40.5 (25.2-56.8)</td>
<td>95.3</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>3.5</td>
<td>5</td>
<td>1,009</td>
<td>PCL(^{100,105}) ARJ at rest(^{91,112}) IT(^{95})</td>
<td>43.2 (28.6-58.4)</td>
<td>94.1</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>411</td>
<td>ARJ at rest and/or PCL(^{69}) PCL or other fixed landmarks(^{97}) ARJ at rest(^{90})</td>
<td>44.6 (25.4-64.6)</td>
<td>92.7</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>6</td>
<td>1</td>
<td>420</td>
<td>PCL(^{88})</td>
<td>61.4 (NA)</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

*Data available in only 18/59 (31%) studies; ARJ: anorectal junction; CI: confidence interval; PCL: pubococcygeal line; IT: ischial tuberosity; NA: not applicable.

**Megarectum**

Prevalence of megarectum in CC was reported in only 2 studies.\(^{79,93}\) Based on findings of barium defecography prospectively performed on 46 HV (28 women), Palit et al.\(^8\) suggested that a rectal diameter of >8.1 cm in men and >6.9 cm in women is indicative of megarectum. Using these parameters, Mohammed et al.\(^{93}\) found a megarectum in 7%
(14/200) of constipated patients. Spazzafumo et al. regarded as abnormal an ampulla >7 cm in diameter on the lateral view, observing this finding in 31% of CC patients.

**Healthy volunteers (HV)**

In only 4 of the 12 controlled studies, the control group was entirely composed of truly HV. A total of 4 studies (2 using barium defecography and 2 MRID) provided normal data by exclusively including >40 subjects.

Despite adopting the same classification system and reporting on a similar gender ratio, the prevalence of intussusception in the study by Shorvon et al. was much higher than that reported by Palit et al. (70% vs 20%, respectively). Rectocele has been much more frequently observed in female (81-100%) than male (0-13%) volunteers on barium defecography.

Goh et al. used MRID to characterize 50 HV (25% females): whilst excessive anorectal junction descent (>3 cm below the pubococcygeal line on maximum strain) was observed in 6% of subjects, prevalence of intussusception, rectocele and enterocele was 0%.

As noted previously, the overlap in presence of structural abnormalities between health and disease is a frequently cited limitation of defecography. Accordingly, grade or severity of abnormality should be considered, and what reflects the pathology (discussed in detail below). Palit et al. proposed that only recto-anal (not recto-rectal) intussusceptions and rectoceles of ≥4.0 cm depth should be considered as truly abnormal findings on barium defecography with regard to size, although it is acknowledged that smaller rectoceles may be clinically relevant in some patients.

Among HV, prevalence of enterocele is rare, ranging from 0% on MRID to only 4% on barium defecography. Extension of the small bowel up to 2 cm below the vaginal apex has been considered as within the normal range.
Pathological structural abnormalities in constipated patients

Prevalence of recto-anal (i.e. Oxford III and IV) intussusception (Figure 4) and external rectal prolapse (i.e. Oxford V) on barium defecography is 23.7% (95%CI, 16.8-31.4; based on 13 studies) and 5.3% (3.1-8.0; based on 16 studies), respectively. When considering large (>4 cm) rectocele only, the prevalence, based on 9 studies, is 15.9% (95%CI, 10.4-22.2) (Figure 4).

**Table 3.** Definitions and pooled prevalence of structural abnormalities on defecography in health and constipation.

<table>
<thead>
<tr>
<th>Structural abnormalities</th>
<th>Health</th>
<th>Constipation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>N studies</td>
</tr>
<tr>
<td></td>
<td>% (95%CI)</td>
<td></td>
</tr>
<tr>
<td><strong>BD</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal prolapse</td>
<td>20-70 (NA)†</td>
<td>2</td>
</tr>
<tr>
<td>External prolapse</td>
<td>0 (NA)</td>
<td>2</td>
</tr>
<tr>
<td>Rectocele</td>
<td>81-100 (NA)§</td>
<td>2</td>
</tr>
<tr>
<td>Enterocele</td>
<td>0 (NA)</td>
<td>2</td>
</tr>
<tr>
<td>Perineal descent</td>
<td>0 (NA)</td>
<td>2</td>
</tr>
<tr>
<td><strong>MRID</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Internal prolapse</td>
<td>0 (NA)</td>
<td>1</td>
</tr>
<tr>
<td>External prolapse</td>
<td>0 (NA)</td>
<td>1</td>
</tr>
<tr>
<td>Rectocele</td>
<td>0 (NA)</td>
<td>1</td>
</tr>
<tr>
<td>Enterocele</td>
<td>0 (NA)</td>
<td>1</td>
</tr>
<tr>
<td>Perineal descent</td>
<td>6 (NA)</td>
<td>1</td>
</tr>
</tbody>
</table>

BD: barium defecography; MRID: magnetic resonance imaging defecography.
* Not seen in health; CI: confidence interval; NA: not applicable; † Oxford I or II (i.e. recto-rectal intussusception); †† Oxford III or IV (i.e. recto-anal intussusception); § <4 cm depth observed in females compared to 0-13% in males (43-56% overall); ‡‡ >4 cm depth.

Functional abnormalities

Patients

On defecography, the diagnosis of a functional abnormality is made using 3 possible features, either combined or in isolation, originally described by Mahieu et al.¹⁹: a) poor opening of the anorectal angle (secondary to poor relaxation or indeed ‘paradoxical’
contraction of the puborectalis muscle); b) poor anal sphincter relaxation; c) incomplete and/or prolonged evacuation based on percentage of contrast expelled and/or time taken, respectively.

Among the 59 studies in patients with CC, diagnostic criteria and prevalence of functional abnormalities were provided in 42 (71%), based on either a) (n = 22); b) (n = 2); c) (n = 2); a+b) (n = 4); a+c) (n = 7); b+c) (n = 1); or a+b+c) (n = 4). Quantitative meta-analysis of these studies, including 4 comparative (barium defecography vs. MRID) studies, showed a pooled prevalence of 24.1% (95% CI, 20.2-28.4) on barium defecography and 25.9 (14.1-39.6) on MRID (Suppl. Figure 6). Prevalence of functional abnormalities was lower in the comparative studies. However, 4 different diagnostic criteria were used.

Prevalence of functional abnormalities in studies where diagnosis was based on the assessment of defecatory dynamics in isolation, compared with those adding parameters of rectal emptying was near identical on barium defecography (23.6% vs. 24.2%, respectively; OR 1.05 [0.93-1.19], p=0.454), but notably different on MRID (23.9% vs. 36.3%; OR 1.81 [95%CI,1.12-2.91], p=0.013). (Table 4).

**Table 4.** Prevalence of functional abnormalities on defecography according to diagnostic criteria in patients with ED.

<table>
<thead>
<tr>
<th>Defecography</th>
<th>Diagnostic criteria</th>
<th>No. studies</th>
<th>No. defecographies</th>
<th>Pooled prevalence (%)</th>
<th>I² (%), p value</th>
<th>OR (95%CI), p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>BD</td>
<td>a±b</td>
<td>26</td>
<td>3,584</td>
<td>23.6 (18.4-29.2)</td>
<td>93.7, &lt;0.001</td>
<td>1.05 (0.93-1.19), 0.454</td>
</tr>
<tr>
<td></td>
<td>a±b+c</td>
<td>9</td>
<td>984</td>
<td>24.2 (18.2-30-8)</td>
<td>79.3, &lt;0.001</td>
<td>1.81 (1.12-2.91), 0.013</td>
</tr>
<tr>
<td>MRID</td>
<td>a±b</td>
<td>4</td>
<td>251</td>
<td>23.9 (6.7-47.1)</td>
<td>93.1, &lt;0.001</td>
<td></td>
</tr>
<tr>
<td></td>
<td>a±b+c</td>
<td>2</td>
<td>230</td>
<td>36.3 (30.2-42.7)</td>
<td>NA</td>
<td></td>
</tr>
</tbody>
</table>

**BD:** barium defecography; **MRID:** magnetic resonance imaging defecography; a: poor opening of the anorectal angle (secondary to non-relaxation or contraction of puborectalis muscle); b: poor anal sphincter relaxation; c: incomplete and/or prolonged evacuation; CI: confidence interval; OR: odds ratio; NA: not applicable.
Table 5 describes the pooled prevalence in studies using either 50-150 ml vs. 151-200 ml vs. >200 ml of contrast, or defecatory desire volume. The volume of rectal contrast used for barium defecography did not influence the prevalence of functional abnormalities, which was slightly lower in studies using up to 150 ml compared with 151-200 ml of rectal contrast (22.0% vs. 27.2%). This finding was borderline statistically significant (OR, 0.85 [95%CI 0.71-1.01]; p=0.064) (Table 5).

**Table 5.** Prevalence of functional abnormalities on barium defecography according to volume of rectal contrast.

<table>
<thead>
<tr>
<th>Rectal contrast (ml)</th>
<th>No. studies</th>
<th>No. defecographies</th>
<th>Pooled prevalence (%, 95%CI)</th>
<th>$I^2$ (%)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>≤150</td>
<td>8</td>
<td>554</td>
<td>22.0 (15.9-28.7)</td>
<td>69.2</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>151-200</td>
<td>9</td>
<td>1,726</td>
<td>27.2 (17.5-38.1)</td>
<td>95.4</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>&gt;200 or DDV</td>
<td>18</td>
<td>2,125</td>
<td>24.7 (19.0-30.8)</td>
<td>89.7</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>

CI: confidence interval; DDV: defecatory desire volume.

Healthy volunteers (HV)

Mahieu et al.\textsuperscript{19} defined criteria for ‘normality’ based on 5 functional parameters in 56 subjects with no symptoms of ED (non-healthy controls): increased anorectal angulation, obliteration of the impression of the puborectalis muscle, wide opening of the anal canal, total evacuation of the rectal contents, and normal resistance of the pelvic floor. In a recent series of 113 asymptomatic women undergoing MRID,\textsuperscript{118} median contrast (ultrasound gel) evacuation was 57%, with 20% proposed as the lower limit of the normal range; this perhaps is a reflection of the supine study position. The authors suggested that only patients who are unable to empty above this cut-off should be considered abnormal on MRID.

Normal findings in constipated patients
Pooled prevalence of normal findings on barium defecography in CC was 16.7% (95% CI, 12.2-21.8) based on 19 studies incorporating a total of 3,086 investigations (Suppl. Figure 7).

**Comparison of barium defecography with MRID in constipated patients**

A total of 5 studies compared barium defecography with MRID. Barium defecography represented the reference standard in all studies, except one adopting the results obtained from the joint analysis of barium defecography and MRID as reference.\(^1\) Vitton et al.\(^9\) compared the accuracy of dynamic anorectal endosonography and MRID with barium defecography as the reference standard in the diagnosis of pelvic floor disorders in 56 women with ED. Diagnostic concordance between barium defecography and MRID did not differ significantly. Concordance rates for MRID were 82% for rectocele, 57% for perineal descent, 93% for enterocele, and 55% for rectal intussusception. Pilkington et al.\(^9\) aimed to establish whether there were measurable differences between barium defecography and MRID in 42 consecutive patients. Anismus (functional dysfunction) was reported in 29% on barium defecography and 43% on MRID. MRID missed 31% of rectal intussusceptions detected on barium defecography. The agreement between grade of rectal intussusception was only fair (k=0.26), with MRID tending to underestimate this. Patients reported that they found it harder to empty their bowel lying in the MRI scanner. Indeed, complete rectal emptying occurred in only 2% of subjects on MRID compared with 29% on barium defecography. This may have negatively impacted MRID sensitivity for detecting rectal intussusception. Zafar et al.\(^1\) reported similar findings in a prospective study of 55 patients with ED undergoing both techniques. Barium defecography detected more rectal intussusceptions than MRID. Again, though not statistically significant, patients achieved higher rates of rectal emptying during barium defecography compared to MRID. Detection rates for rectocele were similar, but barium defecography revealed a
significantly higher number of trapping rectoceles compared to MRID. Furthermore, MRID appeared to underestimate the rectocele size, although it was able to detect a significant number of anatomical abnormalities missed on barium defecography in the anterior and middle pelvic floor compartments. Contrarily, however, higher MRID sensitivities for intussusception have been reported by the 2 most recent comparative studies.\textsuperscript{112, 121} Nevertheless, pooled prevalence of the 5 comparative studies showed that barium defecography was superior to MRID in the detection of intussusception (57.8\% vs. 37.8\%\; OR, 1.52 [95\%CI 1.12-2.14, p=0.009]), although the technique was associated with higher level of embarrassment (qualitatively measured among patients) and/or lower tolerance (54.3\% vs. 30.0\%; OR, 1.73 [95\%CI 1.14-2.62, p=0.008]) (Figure 5).

**DISCUSSION**

This systematic review and meta-analysis demonstrated that the prevalence of structural and functional abnormalities detected by defecography was high, but varied considerably across studies, with high heterogeneity that may reflect variation in measurements, patients or procedural variations. Nevertheless, findings that may be considered truly pathological (i.e. not seen in health) were still frequently observed.

**Interpretation**

Structurally significant (pathological) intussusception (i.e. recto-anal) and rectocele (>4 cm depth) were found in one in four (23.7\%) and one in six (15.9\%) patients with symptoms of CC, respectively (Table 6). Interestingly, their prevalence was higher in prospective than retrospective studies (33.7 [21.0-47.6] vs. 17.1 [10.6-24.7], and 23.1 [14.5-32.9] vs. 11.6 [5.6-19.2], respectively). Despite being adopted by only 1 study,\textsuperscript{103} the Oxford Prolapse Grade system can easily differentiate between an intra-rectal (grade I and II) and intra-anal (grade III and IV) intussusception and is the preferred method to assess prolapse.
severity in patients undergoing corrective surgical procedures. Conversely, poor agreement was found between the 2 studies reporting on outcomes of barium defecography in HV using the classification proposed by Shorvon et al.\textsuperscript{20} despite reporting on a similar gender ratio, the prevalence of intussusception in the study by Shorvon et al.\textsuperscript{20} was much higher than that reported by Palit et al.\textsuperscript{8} (70\% vs 20\%, respectively) and likely reflects the challenge in diagnosing minor clinically insignificant infolding.\textsuperscript{82}

Despite a paucity of information in HV using barium defecography (primarily for ethical reasons), these studies show that prevalence of structural abnormalities in health is not negligible and may lead to over-interpretation of barium defecography, as has already been acknowledged.\textsuperscript{33} Interpreting intra-anal intussusception and large rectocele as truly pathological is in keeping with the findings of Palit et al.,\textsuperscript{8} who showed a rectocele with mean depth of 2.5 ±0.7 cm in 26/28 (93\%) and low grade (recto-rectal) intussusception in 20\% of healthy female volunteers. Similarly, prevalence of >2 cm rectocele, internal, and external prolapse were found in 62\%, 11\%, and 4\%, respectively, in a recent study of 113 healthy females undergoing MRID (published later than final search date, hence not included in this systematic review).\textsuperscript{118}

Prevalence of enterocele ranged from 0\% on MRID\textsuperscript{81} to 4\% on barium defecography.\textsuperscript{20} Given such low prevalence in health, enterocele should be regarded as pathological, and this was found in about one in six CC (16.8\%) patients. Diagnostic yield of barium defecography for enterocele was 6\% higher after oral contrast administration (though did not reach statistical significance); nevertheless, adding this step to conventional barium defecography may be preferable as it allows enterocoele to be categorically confirmed or refuted.

The outcomes of this systematic review support the use of radiology alongside other common tests of ED (i.e. clinical examination, BET, and AM), to enable an accurate
morphological assessment of the posterior pelvic floor compartment. The considerable
disagreement between the results of all current modalities\textsuperscript{111} highlights the need for a
reappraisal of both diagnostic criteria and what represents the ‘gold standard’
investigation. One of the principle challenges will be to promote standardization of the
technique so that results are transferrable between institutions. However, superiority of
defecography over other common tests of ED lies in the detailed assessment of structural
abnormalities, which are frequently observed in patients with CC.

Symptoms of constipation may also affect patients in the absence of any obstructive
structural rectal or pelvic floor features. Spasm/hypertrophy of the puborectalis muscle
was initially proposed as the main pathophysiologic mechanism in this CC subgroup by
Wasserman\textsuperscript{16} in 1964, who reported 4 cases of ‘puborectalis syndrome’, a condition
subsequently named ‘anismus’\textsuperscript{124} or ‘dyssynergia’.\textsuperscript{125, 126} In its broadest sense, the latter
term indicates a failure of recto-anal coordination during straining. Other synonyms
included: ‘spastic pelvic floor syndrome’,\textsuperscript{127} ‘abdomen-levator incoordination’,\textsuperscript{128} ‘immobile
perineum’,\textsuperscript{129} and ‘abdomino-pelvic asynergism’.\textsuperscript{130} More recently, the term ‘functional
defecation disorder’ has been adopted by the Rome classification system to characterize
paradoxical contraction or inadequate relaxation of the pelvic floor muscles and/or
inadequate propulsive forces during attempted defecation.\textsuperscript{2} In this scenario, defecography
may have an important role in the study of recto-anal coordination, especially in light of
recent evidence discrediting the diagnostic accuracy of anorectal manometry for
dyssynergic defecation.\textsuperscript{131}

Finally, further studies should clarify whether patient position (supine in all included studies
on MRID) may explain the increased sensitivity of barium defecography over MRID in the
detection of intussusception, being the former associated with higher rates of complete or
nearly complete rectal emptying. Until new evidence is available, barium defecography may be considered when in doubt of intussusception on MRID.

**Limitations**

We must acknowledge that our pragmatic threshold of including studies reporting outcomes on more than 40 subjects (Figure 1) served to exclude small case series that often reported on early experience with the techniques, but also left out a significant number of studies from reputed and established institutions with high-quality research. Furthermore, potential sources of heterogeneity were not explored.

Except for comparative (barium defecography vs. MRID) studies, OR calculations were not derived from the same studies. Hence, a significant OR value does not indicate diagnostic accuracy between the two tests.

In conclusion, pathological structural abnormalities, as well as functional abnormalities, are common in CC patients. Since structural abnormalities cannot be evaluated using non-imaging test modalities (balloon expulsion and anorectal manometry) defecography should be considered first-line diagnostic test, if resources allow.

**Table 6.** Pooled prevalence for specific findings according to test.

<table>
<thead>
<tr>
<th>Test findings</th>
<th>Anorectal manometry</th>
<th>Balloon expulsion test</th>
<th>Defecography</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>52%</td>
<td>60%</td>
<td>17%</td>
</tr>
<tr>
<td>Abnormal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Functional</td>
<td>48%*</td>
<td>40%‡</td>
<td>83%</td>
</tr>
<tr>
<td>Structural</td>
<td>100%</td>
<td>100%</td>
<td>24%</td>
</tr>
<tr>
<td>Intussusception</td>
<td>NA</td>
<td>NA</td>
<td>29%†</td>
</tr>
<tr>
<td>Recto-anal (Oxford III-IV)</td>
<td>NA</td>
<td>NA</td>
<td>24%†</td>
</tr>
<tr>
<td>External prolapse (Oxford V)</td>
<td>NA</td>
<td>NA</td>
<td>5%†</td>
</tr>
<tr>
<td>Rectocele &gt;4 cm</td>
<td>NA</td>
<td>NA</td>
<td>16%†</td>
</tr>
<tr>
<td>Enterocele</td>
<td>NA</td>
<td>NA</td>
<td>17%†</td>
</tr>
<tr>
<td>Megarectum</td>
<td>NA</td>
<td>NA</td>
<td>7%†</td>
</tr>
</tbody>
</table>
Dynamic perineal descent  NA  NA  45%†

†Data from a previous meta-analysis. NA: not applicable. †Truly pathological abnormalities (i.e. not seen in health). *Defined as dyssynergic pattern (i.e. paradoxical contraction or inadequate relaxation of the anal sphincter on attempted defecation).
‡Defined as patients unable to expel the balloon after 5 min seated on a commode.

FIGURE LEGENDS

Figure 1: PRISMA diagram.

Figure 2: Oxford Grading System for rectal prolapse.

Figure 3: Forest plot showing rates of enterocele on barium defecography in studies with or without the use of oral contrast (percentage of patients). KEY: ES = effect size; CI = confidence interval.

Figure 4: Forest plot showing rates of structurally significant intussusception (Oxford III and IV; A) and rectocele (≥4 cm depth; B) on barium defecography (percentage of patients). KEY: ES = effect size; CI = confidence interval.

Figure 5: Pooled prevalence and 95% confidence interval (CI) for specific findings according to radiologic technique. KEY: BD = barium defecography; MRID = magnetic resonance imaging defecography; OR = odds ratio.

Suppl. Figure 1: Component 1 (10-criterion checklist) for quality assessment of studies.

KEY: Green = Yes; Orange = Partial or Unclear; Red = No.

Suppl. Figure 2: Forest plot showing rates of internal rectal prolapse on barium defecography (percentage of patients).

Suppl. Figure 3: Forest plot showing rates of internal rectal prolapse on magnetic resonance imaging defecography (percentage of patients).

Suppl. Figure 4: Forest plot showing rates of rectocele on barium defecography (percentage of patients).
Suppl. Figure 5: Forest plot showing rates of perineal descent on barium (A) and magnetic resonance imaging (B) defecography (percentage of patients).

Suppl. Figure 6: Forest plot showing rates of functional abnormalities on barium (A) and magnetic resonance imaging (B) defecography (percentage of patients).

Suppl. Figure 7: Forest plot showing rates of normal findings on barium defecography (percentage of patients).
ACKNOWLEDGEMENTS

The authors would like to thank Caterina Larosa for her contribution in the illustration of the Oxford Grading System for rectal prolapse.

AUTHOR CONTRIBUTIONS

UG and SMS conceived the study and developed the first draft. UG and HH handled literature search and data extraction. GLDT conducted the meta-analyses. UG, HH, GLDT, SAT, CHK, and SMS contributed in the processes of defining criteria for inclusion and exclusion of studies, reviewing and assessing the primary studies, discussing findings, drawing conclusions, as well as the completion of the manuscript. All authors read and approved the final manuscript.

COMPETING INTERESTS

None declared.

FUNDING

None.

REFERENCES

1. Suares NC, Ford AC. Prevalence of, and risk factors for, chronic idiopathic constipation in the community: systematic review and meta-analysis. Am J Gastroenterol 2011;106(9):1582-91; quiz 81, 92. doi: 10.1038/ajg.2011.164


96. Regadas FS, Haas EM, Abbas MA, et al. Prospective multicenter trial comparing echodefecography with defecography in the assessment of anorectal dysfunction in patients with obstructed


36


