

Surgery for constipation: systematic review and clinical guidance

Paper 1: Introduction & Methods

C. H. Knowles*, U. Grossi*, E. J. Horrocks*, D. Pares†, P. F. Vollebregt*, M. Chapman‡, S. R. Brown§, M. Mercer-Jones¶, A. B. Williams**, R. J. Hooper††, N. Stevens††, J. Mason‡‡, on behalf of the NIHR CapaCiTY working group§§ and The Pelvic floor Society¶¶

*National Bowel Research Centre, Blizard Institute, Queen Mary University London, London, UK, †Hospital Germans Triás i Pujol, Barcelona, Spain, ‡Good Hope Hospital, Heart of England NHS Trust, Birmingham, UK, §Sheffield Teaching Hospitals, Sheffield, UK, ¶Queen Elizabeth Hospital, Gateshead NHS Trust, Gateshead, UK, **Guy's and St Thomas' NHS Foundation Trust, London, UK, ††Pragmatic Clinical Trials Unit, Blizard Institute, Queen Mary University of London, London, UK, ‡‡University of Warwick, Coventry, UK, §§National Institute for Health Research: Chronic Constipation Treatment Pathway, UK, and ¶¶Affiliate section of the Association of Coloproctology of Great Britain and Ireland, London, UK

Abstract

Aim This manuscript provides the introduction and detailed methodology used in subsequent reviews to assess the outcomes of surgical interventions with the primary intent of treating chronic constipation in adults and to develop recommendations for practice.

Method PRISMA guidance was adhered to throughout. A literature search was performed in public databases between January 1960 and February 2016. Studies that fulfilled strictly-defined PICOS (patients, interventions, controls, outcome, and study design) criteria were included. The process involved two groups of participants: (i): 'a clinical guidance group' of 18 UK experts (including junior support) who performed the systematic reviews and produced summary evidence statements (SES) based strictly on data synthesis in each review. The same group then produced prototype graded practice recommendations (GPRs) based on coalescence of SES and expert opinion; (ii): a European Consensus group of 18 ESCP (European Society of Coloproctology) nominated experts from nine European countries evaluated the appropriateness of each prototype GPR based on published RAND/UCLA methodology.

Results An overview of the search results is provided in this manuscript. A total of 156 studies from 307 full text articles (from 2551 initially screened records) were

included, providing data on procedures characterized by: (i) colonic resection ($n = 40$); (ii) rectal suspension ($n = 18$); (iii) rectal wall excision ($n = 44$); (iv) recto-vaginal septum reinforcement ($n = 47$); (v) sacral nerve stimulation ($n = 7$). The overall quality of evidence was poor with 113/156 (72.4%) studies providing only Oxford level IV evidence. The best evidence was extracted for rectal excisional procedures, where the majority of studies were Oxford level I or II. The five subsequent reviews provide a total of 99 SES (reflecting perioperative variables, efficacy, harms and prognostic variables) that contributed to 100 prototype GPRs covering patient selection, procedural considerations and patient counselling. The final manuscript details the 85/100 GPRs that were deemed appropriate by European Consensus (remaining 15 were all uncertain) and future research recommendations.

Conclusion This manuscript and the following 6 papers suggest that the evidence base for surgical management of chronic constipation is currently poor although some expert consensus exists on best practice. Further studies are required to inform future commissioning of treatments and of research funding.

Keywords Constipation, surgery, obstructed defaecation

Introduction

Constipation is common in adults and children with up to 20% of the population reporting symptoms

Correspondence to: Prof. Charles H. Knowles, NBRC, 1st Floor Abernethy Building, 2 Newark St, London E1 2AT, UK.
E-mail: ch.knowles@qmul.ac.uk

depending on the definition used (2–28% adults; 0.7–30% children) [1–3]. Chronic constipation (CC), usually defined as more than 6 months of symptoms, is less common but results in 0.5 million UK GP consultations per annum. A proportion of the population suffer symptoms that are both chronic and more disabling (probably about 0.4% population) [4]. Such patients,

who are predominantly female [5], are usually referred to secondary care with many progressing to tertiary specialist investigation. Patient dissatisfaction is high in this group; nearly 80% feel that laxative therapy is unsatisfactory [6] and the effect of symptoms on measured QOL is significant [7]. CC consumes significant healthcare resources. In the US in 2012, a primary complaint of constipation was responsible for 3.2 million physician visits resulting in (direct and indirect) costs of \$1.7 billion [8]. In the UK, it is estimated 10% of district nursing time is spent on bowel control [9] and the annual spend on laxatives exceeds £117 m, with 18.3 million prescriptions in 2014 of which 91% were for stimulant and osmotic laxatives (Health and Social Care Information Centre) [10].

The act of defaecation is dependent on the coordinated functions of the colon, rectum and anus. Considering the complexity of neuromuscular (sensory and motor) functions required to achieve planned, conscious, and effective defaecation [11], it is no surprise that disturbances to perceived 'normal' function occur commonly at all stages of life. Clinically, such problems commonly lead to symptoms of obstructed defaecation e.g. straining; incomplete, unsuccessful or painful evacuation; bowel infrequency; abdominal pain and bloating. After exclusion of secondary causes (obstructing colonic lesions, neurological, metabolic and endocrine disorders), the pathophysiology of CC can broadly be divided into problems of colonic contractile activity (and thus stool transit) and problems allied to rectal emptying (evacuation disorder). A combination of clinical expertise and specialist radio-physiological investigations can determine which patients have slow colonic transit, evacuation disorder, both (in whom transit is usually characterized by a left-sided delay) or neither (no abnormality found with current tests) [12]. Evacuation disorders can be further subdivided into those with a structurally significant pelvic floor abnormality (usually as a consequence of pelvic floor weakness or injury) e.g. rectocele or internal prolapse (intussusception), and those characterized by a dynamic failure of evacuation without structural abnormality: most commonly termed 'functional defaecation disorder (FDD) [13]' (Fig. 1).

The management of CC is a major problem due to its high prevalence and lack of widespread specialist expertise. In general, a step-wise approach is undertaken, with first line conservative treatment such as lifestyle advice and laxatives (primary care) followed by nurse-led bowel re-training programs, sometimes including focused biofeedback and psychosocial support (secondary/tertiary care). Although these treatments may improve symptoms in more than half of patients [14], patients with intractable symptoms and impaired

QOL may subsequently be offered a range of surgical interventions.

Surgical decision-making is greatly influenced by local expertise, commissioning or reimbursement, and personal enthusiasm for particular interventions. While robust diagnosis of specific pathophysiologies combined with multidisciplinary team discussion may help direct surgery, in the absence of an agreed pathway to stratify patients, there is a current large and difficult-to-justify variation in surgical practice that continues to risk inadequately-informed and potentially harmful interventions being offered. The need to reduce such variations in practice, based on available evidence, has been a recurrent theme of recent national specialty group discussions (e.g. ACPGBI) with various initiatives proposed. As part of the Chronic Constipation Treatment Pathway (CapaCiTY) programme funded by National Institute of Health Research (NIHR), a multi-disciplinary working group was convened in July 2014 to address this need. This group of medical and nursing experts included members of The Pelvic Floor Society and urogynaecology expertise derived from the International Continence Society (ICS). As a prelude to developing new evidence from trials within the CapaCiTY programme, it was agreed that the current surgical evidence base would benefit from coalescence in the form of systematic review and graded practice recommendations. This paper and the accompanying subsequent six papers address this aim.

Methods

Systematic review

Protocol and registration

The authors developed the protocol for review, detailing pre-specified methods of the analysis and eligibility for the review in accord with 2009 PRISMA guidance [15] using also the new reporting elements derived from the 2016 harms checklist [16]. While the protocol was not registered, a description of the NIHR CapaCiTY programme is available in the public domain (<http://www.isrctn.com/ISRCTN11747152>) and has been presented nationally (DDF meeting, London 2015; National Pelvic Floor Meeting, Manchester 2015).

Eligibility criteria

Study characteristics

Study characteristics were defined using the PICOS framework. Search term definitions were inclusive, promoting a sensitive search of studies reporting surgical interventions for chronic constipation.

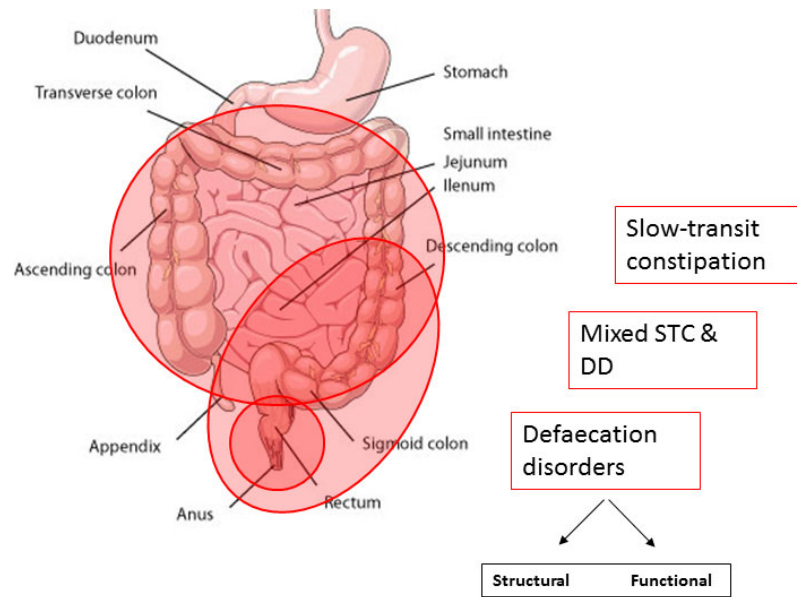


Figure 1 Schematic of pathophysiology of chronic constipation.

Population: The review aimed to identify studies of patients undergoing surgical interventions with the primary intent of treating chronic constipation. The definition of chronic constipation is neither straightforward nor uniformly applied [17]. On this basis, all common terms encompassing problematic defecation were used (see search strategy syntax: Appendix I). However, several pelvic floor procedures may be performed commonly for non-chronic constipation indications. Examples include pelvic organ prolapse syndromes where the physical prolapse or other organ dysfunctions of the vagina or bladder are the main motivation for surgery. While such patients invariably also have some degree of defaecatory problems, and their perioperative data could still be used to inform procedural safety, these patients may phenotypically differ at baseline and in response to surgical intervention even if the intervention itself is identical or at least similar. Cochrane reviews such as ‘surgical repair of pelvic organ prolapse in women’ [18] and of surgical management of external rectal prolapse [19] include some RCTs where defaecatory symptoms are recorded as a secondary outcome or as a complication but not as a primary presenting complaint of the population studied. Thus these were ineligible for inclusion. Similarly, for colonic excisional procedures, patients with the very rare diagnoses of adult Hirschsprung disease or idiopathic megacolon-megarectum [20] were considered distinct from chronic constipation and thus not included. Some studies reported outcomes on two populations, only one of which was eligible e.g. internal and external rectal prolapse. Where such data could not be separated by

population, the study was also deemed ineligible for inclusion.

A minimum population sample of 20 patients was imposed for eligibility. This threshold was taken to exclude case reports and small case series that often reported a single surgeon’s personal experience or early experience of experimental procedures.

Intervention: Surgical procedures for chronic constipation are subject to heterogeneous descriptions. On this basis, an iterative approach was taken by cross referencing e.g. with textbook reference lists to ensure that all terms in common usage were incorporated in the eventual search strategy. These included some genuine procedural variations but also multiple small changes in syntax for the same procedure e.g. ‘stapled transanal rectal resection’ vs ‘stapled transanal rectum resection’. A decision was taken by the review team that results would be grouped by five main approaches to surgically treating chronic constipation: (i) colonic resection, (ii) hitching procedures of the rectum (rectal suspension); (iii) excisional procedures of the rectal wall (rectal excision); (iv) reinforcement of the rectovaginal septum (RV reinforcement); and (v) sacral nerve stimulation (SNS). This approach was taken because initial review (Oct 2014) determined that other procedures either lacked sufficient evidence for review. The first major exclusion on this basis were stomas leading to intestinal discontinuity or for the purpose of administering bowel irrigation (continence enema). It is acknowledged that in the real world many patients have stomas either deliberately or as an eventual outcome of other surgery.

However, eligible studies were sparse after application of inclusion criteria and markedly heterogeneous between and within studies (patients and techniques). Other procedures were excluded if still considered experimental e.g. colonic exclusion procedures [21].

Comparisons: Studies were eligible regardless of whether they were retrospective or prospective in design, controlled or uncontrolled. Only a minority of studies reported more than one procedure or more than one population.

Outcomes: Studies were broadly eligible if they provided extractable data on benefit (treatment efficacy), risk (harms) or both. For efficacy, inclusion necessitated the acceptance of the huge disparity in quality of outcomes reporting that are well acknowledged in the literature [14], with a heavy reliance on estimates of global patient satisfaction with the procedure (an indirect measure of the patients own judgement of their post-operative state compared to their pre-operative state). Studies of physiological and anatomical outcomes alone were excluded since these are generally regarded as a poor surrogate of efficacy in this patient population [22]. Because the outcomes of surgical interventions for chronic constipation are known to exhibit a ‘honeymoon period’ in the months immediately following surgery, a minimum (mean or median) follow up of 12 months was applied for eligibility. It is acknowledged that enforcement of this criteria excluded some level I studies. Several studies reported the outcomes of more than one procedure. Where such data could not be separated by procedure, these were not included (often resulting in study ineligibility).

Report characteristics

Year of publication: Any publication date was eligible as covered by database search from 1960 to the date of final search (22nd February 2016).

Language: Due to the large number of studies retrieved, it was decided to include only studies with full text in the English language. While the numbers of foreign language studies were small, these have been detailed for the reader in ‘reasons for exclusion’ at the full-text stage (rather than at the abstract screening stage). There is reasonable evidence to suggest that searching only in English does not have an adverse effect on the quality of systematic reviews [23].

Type of study: Only peer-reviewed publications reporting primary data were eligible. Thus reviews, editorials, letters and other forms of secondary expert opinion were excluded at the screening stage. Only full

manuscripts were eligible thus conference abstracts and proceedings were also excluded. No constraint was imposed based on level of evidence. This decision was taken in the knowledge that the vast majority of data would be extracted from case series rather than higher quality study types.

Information sources

The senior author (CK) performed a comprehensive search of the literature on 22nd February 2016 using PubMed and Evidence Based Medicine reviews (including the Cochrane database of systematic reviews and the Cochrane central register of controlled trials). A preliminary search in 2014 had determined that Embase and Web of Science led to almost 2000 duplicate records with no additional yield. Search terms used a sensitive combination of population, intervention and report terms. A keyword and hand search was used within relevant Cochrane systematic reviews. The specific search terms are listed in Appendix I.

Study selection

Screening was performed at the abstract level by the senior author (CK), excluding studies not meeting eligibility criteria where this could be readily determined from the abstract alone. Full-text copies of all remaining English language studies were obtained and assessed by reviewers, who were un-blinded to the names of studies, authors, institutions or publications. Disagreement regarding inclusion was resolved by the senior author (CK). Duplicate data sets generated from the same cohort of patients were excluded with the larger population size and longer follow-up cohort included at the expense of earlier reports from the same cohort. In instances of doubt, authors from the relevant institutions were contacted to confirm or refute any repetition of results (performed on three occasions).

Search results were cross-referenced to bibliographies from other sources (previous reviews and book chapters). Care was taken that any studies missed by the original search met the strict inclusion criteria and did not circumnavigate the carefully-defined search strategy especially in relation to population terms.

Data collection process

Outcome data were extracted by the junior authorship team (UG, EJH, DP, PFV) paired with one senior author for each procedure: colonic resections (CK); rectal hitching procedures (SB); rectovaginal septum reinforcement (ABW); rectal wall excision (MM-J); sacral nerve stimulation (SP). Data were extracted to a standardized template (Microsoft Excel spreadsheet) including study characteristics and outcome data (see below).

For each procedure, one reviewer extracted the data and one verified content.

Data items

A full list of data fields is included in Table 1 (with annotation). These followed the PICOS framework with outcomes broadly divided into those assessing harms (intra- and perioperative complications and long-term adverse outcomes), and those assessing efficacy: global success ratings and functional outcomes (organized into validated symptom, QOL scoring instruments and individual symptoms). For perioperative complications, some consideration was given to classifying complications by established systems e.g. Clavien-Dindo however inconsistencies in reporting made this unfeasible. Data were not collected in relation to cost effectiveness which was deemed to fall outside the remit of the process aims. To simplify data extraction and presentation, for ordinal data, summary statistics were extracted as mean or median (with SD when provided).

Individual study quality and risk of bias

The methodological quality of all individual included studies was assessed by the senior author (CK) and classified in accord with Oxford CEBM levels of evidence definitions for 'therapy or harm' [16]. The following rules were applied accepting that distinguishing study designs can be problematic for observational studies [24]:

- 1 A study was deemed prospective if this was categorically stated or if patients were 'enrolled' or 'recruited' to a study that systematically recorded pre- and post-operative data. All other studies were assumed to be retrospective.
- 2 A cohort study was defined as one designed to address a clear stated aim or hypothesis using specified analytical methods. In general, these included a comparison group related either to the relative efficacy of more than one specified procedure or to patient selection where a specified baseline 'risk factor' was analysed in relation to relative success or failure of the intervention.
- 3 A case series was defined as a report of observations based on clinical practice. Such studies may generate hypotheses by *post-hoc* case comparisons.
- 4 For randomized trials and cohort studies, Cochrane risk of bias tools were applied [<http://ohg.cochrane.org/sites/ohg.cochrane.org/files/uploads/Risk%20of%20bias%20assessment%20tool.pdf>] and used to distinguish between high and low quality RCTs (Oxford level 1b or level 2b) and high and low quality cohort studies (Oxford level 2b or level 4). Case control studies were assessed using the National Institutes of Health (NIH) quality assessment tool [<http://www.nhlbi.nih.gov/health-pro/guidelines/>

in-develop/cardiovascular-risk-reduction/tools/case-control] (Oxford level 3b or 4).

Further sub-analysis of the quality of observational studies (e.g. compliance with STROBE, Newcastle-Ottawa or MINORS) was not undertaken as it was felt that this would add little to the overall assessment of quality.

Summary measures

Results were tabulated by outcome and described with appropriate summary statistics (percentages, means and ranges). For very rare events, the aggregate number and denominator were reported. Quantitative data synthesis was performed for key outcomes using meta-analysis in STATA SE v14. Pooled proportions and means were estimated, permitting exploration of heterogeneity and bias. Where continuous measures failed to report measures of variance these were approximated as range/4. Random effect meta-analytic models were estimated to characterise rates of events and heterogeneity between studies, with sub-grouping by procedure. Where studies did not provide data in a useful summary form, available data were tabulated but not included in the meta-analysis. Results were presented as aggregate means with confidence intervals and graphically displayed within Forest plots. For pooled studies, the I^2 value (reflecting intra-group heterogeneity) was reported and interpreted in accord with published guidance where 0–40% = heterogeneity might not be important, 30–60% = moderate heterogeneity, 50–70% substantial heterogeneity and 75–100% = considerable heterogeneity [25]. The magnitude and direction of effect, and strength of evidence P -value from the chi-squared test, were used to interpret the importance of heterogeneity.

Evidence within reviews was predominantly provided by observational cohort data with relatively few experimental studies (trials) identified. Consequently, the reviews analyse all studies as individual cohorts, by procedure, to achieve inclusion and consistency; pooled findings are compared with the findings of individual trials. Where several trials were identified within a review (e.g. rectal excision procedures) meta-analyses was performed with sub-grouping by procedure and by evidence grade. Findings by evidence grade were reported only when they deviated qualitatively from the overall pooled summary. Given the nature and reporting of data, study-level meta-regression was not attempted.

Risk of bias across studies

Publication bias was assessed for outcomes where meta-analysis was performed. Other limited analysis was performed based on study size, design and publication date where this contributed to interpretation. Subgroup analysis was explored for the main procedural variations.

Table 1 Data field for systematic review.

| Data extract | Description | Notes |
|------------------------------|-----------------------|---|
| Study characteristics | | |
| First author | Text ^(num) | With citation number |
| Year publication | Text | To 2016 |
| Number of pts | Number | Ordinal integer |
| Follow up | Months | Mean or median as documented in study (integer) |
| Study design | Text abbreviation | As Oxford CEBM levels of evidence |
| Evidence grade | IA–IV | As Oxford CEBM levels of evidence |
| Population | | |
| Disease | Text abbreviation | As supplied key |
| Sex ratio | Female:male | Ratio not simplified |
| Age | Years (integer) | Mean or median as supplied (range) |
| Intervention | | |
| Operation (s) | Text abbreviation | As supplied key |
| Op duration | Minutes | Mean (integer) |
| Length of stay | Days | Mean to 1 decimal place |
| Outcomes | | |
| Harms | | |
| Perioperative | | |
| Total cx | Percentage | % to 1 decimal place |
| Infective cx | Percentage | % to 1 decimal place |
| Bleeding cx | Percentage | % to 1 decimal place |
| Proc specific cx | Percentage | As per specific procedure: % to 1 decimal place |
| Mortality | Percentage | % to 1 decimal place |
| Repeat intervention | Percentage | Generally procedure specific for complications or poor functional outcome |
| Mortality rate | Number | Absolute number over reported denominator |
| Adverse long-term symptoms | Percentage | Some procedural specificity: Includes re-operation rate where relevant |
| Efficacy | | |
| Global success rating | Scale | Very commonly employed: % patients with good or excellent outcomes unless specified |
| Symptom scores | Count | Several variably validated summative scoring instruments: pre and post or post only as available: mean + SD |
| Individual symptoms | Percentage | Some procedural specificity: pre and post or post only as available |
| QOL measures | Count or scale | Few instruments used: pre and post or post only as available: mean + SD |

Cx, complications.

Funding statement

This review was funded by the NIHR Programme Grant for Applied Research Programme (Ref: RP-PG-0612-20001) who had no role in study design, data collection, analysis, interpretation, writing of the report or the decision to submit the paper.

Clinical guidance development

Aims

The process had three main aims:

1 Development of summary evidence statements;

2 Development of graded practice recommendations;

3 Development of summary research recommendations.

Development of summary evidence statements

Summary evidence statements were produced by the *Clinical Guidance Group (CGG)*. This group was convened in summer 2014. A final list of participants was selected primarily from colorectal surgeons, gastroenterologists, urogynaecologists and specialist nurses with a strong interest in functional colorectal and pelvic floor disorders. This group included all senior authors of the

five reviews and associated junior investigators. Methodological expertise was provided by Professor James Mason (University of Warwick), and NHS Specialised Services stakeholder representation by Mr Mark Chapman. A series of meetings followed (Bristol, November 2014; London, June 2015; Manchester, November 2015; and Edinburgh, July 2016) at which the evolving summary evidence statements (from reviews) were eventually ratified and prototype clinical practice recommendations drafted.

The CGG used 'focus group' methodology to gain consensus by *in silico* and face to face meetings. The number of participants (> 12), and four rounds of written revisions fulfilled the basic criteria required for a guideline decision group (National Institute for Health and Clinical Excellence, April 2007) and allowed a sufficiently reliable process at an acceptable cost in terms of travel, expenses etc. The heterogeneity of the group (specialty, nationality, expertise) was deemed desirable to be representative of a range of stakeholders. Agreement was defined without 'weighting' of any participant's views, although some participants contributed more than others to the process.

Using the synthesis of the evidence base the group drafted statements of evidence based on best evidence available (which varied significantly by procedure). The clinical guidance group discussed, revised and graded summary statements of evidence level using the Oxford 2009 CEBM system (<http://www.cebm.net/oxford-centre-evidence-based-medicine-levels-evidence-march-2009>) (Table 2) based on the review of evidence. For clarity, roman numerals (I–IV) were used to denote summary levels of evidence for graded evidence in contrast to Arabic numerals for individual studies, e.g. 1a, 2b etc. Summary levels could apply either positively or negatively to each procedure. Care was taken to avoid any contamination of expert opinion into statements, these thus solely reflecting summated evidence from systematic review. Some language used in summary evidence statements was deliberately chosen to reflect use of pooled data. Thus the term 'typical' or 'typically' specifically denotes that data for the event in question have been derived from random effects analysis.

Development of graded practice recommendations (GPRs)

This had two main stages: (i) development of 'prototype' GPRs by the *Clinical Guidance Group*, and (ii) development of a final GPR list by a *European Consensus group*. This approach, including the methodology used (RAND/UCLA – see below) is established and has been used previously in the coloproctology field [26].

Development of prototype GPRs: After a common understanding of the evidence was established, group

discussion balanced clinical experience and evidence summaries to arrive at shared judgements about recommendations for care, thus deriving relevant recommendations for decision making in clinical practice. Group processes risk personal bias based on 'eminence' or 'eloquence' if led and supported ineffectively: adequate methodological support in the use of evidence and dialectic was provided to support the process to ensure a balance of views as well as to promote generalizability and impact. This stage embodied summary evidence statements (from each review), data from some excluded level I studies (e.g. RCTs that were excluded for short follow up or published after the review date) (a further search was run by CK on 03.10.16 for the date range 22.02.16 to 03.10.16 including original terms and 'clinical trial') and expert opinion derived from the decision group and selected prior published guidance documents (Oxford 5) (Fig. 2).

Final grading followed Oxford CEBM recommendations (A–D) [27] [Table 3]. As with levels of evidence the grades of evidence could apply either positively or negatively to the procedure.

Development of final GPRs: The European Consensus group comprised a panel of European experts (colorectal and pelvic floor surgeons) nominated by the European Society of Coloproctology (ESCP). Twenty experts were invited from 10 European countries of whom 18 participated from nine countries (Appendix II).

Consensus methodology was derived from the RAND/UCLA Appropriateness Method (Prepared for Directorate General XII, European Commission 2001) [28]. Prototype Graded Practice Recommendations (derived from the clinical guidance group) were presented (on a spreadsheet) for each procedure under three subheadings: 'patient selection', 'procedural considerations' and 'patient counselling'. For each, a number of GPRs were listed, each with associated levels of evidence and grade of prototype recommendation. For each, consensus panellists were asked 'Does this recommendation lead to an expected health benefit that exceeds the expected negative consequences of its introduction?' Examples of health benefits in this context could be improved surgical outcome, improved patient experience, improved functional capacity etc.; the negative consequences could include increased morbidity, anxiety, pain, time lost from work, denial of an investigation or treatment. Panellists were asked to base their judgement on clinical grounds only, i.e. exclusive of financial cost [29].

Responses to each listed recommendation used a linear analogue scale of 1–9 to assess views on the benefit-to-harm ratio. Using this scale, a score of 1–3 indicated that they expected the harms of introducing the recommendation to greatly outweigh the expected benefits

Table 2 (a) Oxford CEBM (2009) summary levels of evidence and (b) grades of recommendation.

| (a) Summary level of evidence | Type of studies | Evidence included specific to review | Notes specific to review exclusions |
|-------------------------------|--|--------------------------------------|---|
| I | High quality RCT All or none study | Oxford 1b, 1c* | 1a (SR RCTs) excluded since no secondary research included in systematic review |
| II | Poor quality RCT Individual high quality cohort study Ecological study | Oxford 2b 2c* | 2a (SR cohort studies) excluded since no secondary research included in systematic review |
| III | Individual high quality case-control study | Oxford 3b* | 3a (SR case-control studies) excluded since no secondary research included in systematic review |
| IV | Case series and poor quality cohort and case-control studies | Oxford 4 | The majority of studies included in systematic review |
| V | Expert opinion, bench research | Oxford 5 | Excluded in systematic review |

| (b) Grades of recommendation | Evidence required |
|------------------------------|--|
| A | Consistent level 1 studies |
| B | Consistent level 2 or 3 studies or extrapolations from level 1 studies |
| C | Level 4 studies or extrapolations from level 2 or 3 studies |
| D | Troublingly inconsistent or inconclusive studies of any level (I–IV) |
| N | Recommendation based on clinical understanding in the absence of evidence [†] |

*No studies of these designs found by search for any procedure.

†But where a recommendation was considered necessary to highlight the absence of evidence for an important practice point.

and a score of 7–9 that the expected benefits greatly outweighed the expected harms. A middle rating of 4–6 could mean either that the harms and benefits were considered about equal or that the panellist was unable to make a judgement for the recommendation. Panellists were asked to try and provide a response for all listed recommendations.

Responses were analysed in accordance with the first phase of RAND/UCLA guidance, with each recommendation classified as ‘appropriate,’ ‘uncertain’ or ‘inappropriate’ according to the panellists’ median score and the level of disagreement. Indications with median scores in the 1–3 range were classified as inappropriate, those in the 4–6 range as uncertain, and those in the 7–9 range as appropriate. All indications rated ‘with disagreement,’ whatever the median, were classified as uncertain. ‘Disagreement’ here basically implied a lack of consensus, either because of polarisation or spread over the entire scale (defined for a sample of 18 panellists as > 5 rating the indication outside the 3-point region [1–9,28]). Further phases of consensus following discussion to reduce variation were not conducted.

Summary research recommendations

One of the initial drivers for this process (NIHR CapaCiTY) was the need to define the main evidence needed for future surgical trials of patients with CC. During the development of this guidance, some trials have commenced patient recruitment such as CapaCiTY study 3 (RCT of laparoscopic ventral rectopexy). There is however still a great need to define research questions that could inform future UK and international commissioning of research funding. Research recommendations

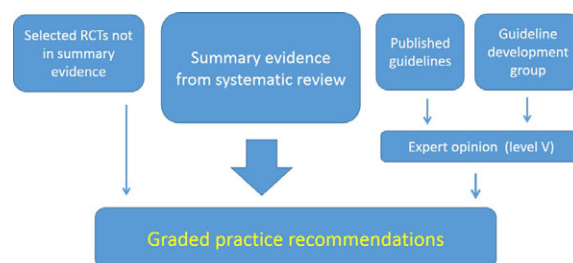


Figure 2 Schematic showing process of transition from summary evidence statements to graded practice recommendations.

have been attributed a priority (high, medium or low) based on the expert opinion of the current working group and may help inform discussion about future funding priorities.

Presentation of results

In view of the large scale of the systematic review and prototype guidance process, results have been presented as a series of separate manuscripts:

- 1 Overview of search results and study characteristics (this manuscript);
- 2 Systematic review results and summary evidence statements for colonic resection;
- 3 Systematic review results and summary evidence statements for procedures characterized by rectal suspension;
- 4 Systematic review results and summary evidence statements for procedures characterized by rectal wall excision;
- 5 Systematic review results and summary evidence statements for procedures characterized by rectovaginal septum reinforcement;
- 6 Systematic review results and summary evidence statements for sacral nerve stimulation;
- 7 Coalescence of systematic review data, summary of graded practice recommendations and research recommendations.

Table 3 Reviewed studies by main procedure type and evidence level.

| Procedure | Number of reviewed studies by evidence level | | | | |
|------------------------------|--|----|----|-----|-------|
| | 1b | 2b | 3b | 4 | Total |
| Colonic resection | 0 | 1 | 0 | 39 | 40 |
| Rectal suspension procedures | 0 | 2 | 0 | 16 | 18 |
| Rectal excisional procedures | 3 | 26 | 0 | 18 | 47 |
| RV Reinforcement procedures | 1 | 10 | 0 | 33 | 44 |
| Sacral nerve stimulation | 0 | 0 | 0 | 7 | 7 |
| ALL | 4 | 39 | 0 | 113 | 156 |

The main conclusions of this process were presented at the Pelvic Floor Society Meeting in Cardiff, January 2017.

Overview of search results

Study selection

Figure 3 (PRISMA flow diagram) shows the results of population and intervention term searches with reasons for exclusion of studies at the full text review stage.

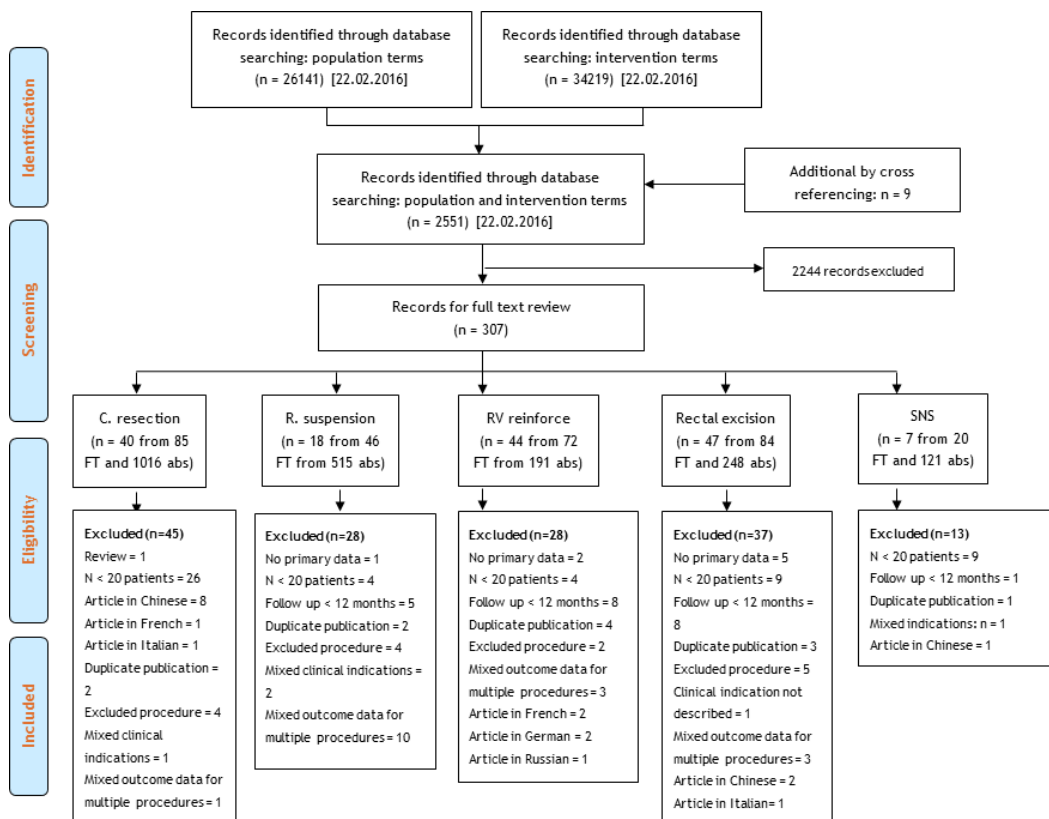


Figure 3 PRISMA diagram showing all review results.

Study characteristics

Table 3 gives information on the overall study characteristics and by procedure. Detailed data on individual reviewed studies are provided by procedure type in the accompanying papers. It can readily be noted that the overall quality of evidence was poor with 113/156 (72.4%) providing only level IV evidence. The best evidence to date exists for rectal excisional procedures where the majority of studies were level I or II. This is discussed further in the final graded practice recommendations and research recommendations paper.

Acknowledgements

We would like to acknowledge the following people in relation to the NIHR CapaCiTY working group: Mr Steven Brown, Sheffield Teaching Hospitals NHS Foundation Trust; Mr Kenneth Campbell, The University of Dundee, NHS Education for Scotland; Mr Mark Chapman, Heart of England NHS Trust; Mr Andrew Clarke, Poole Hospital NHS Foundation Trust; Mr Neil Cruickshank, Sandwell and West Birmingham Hospitals NHS Trust; Mr Anthony Dixon, University of Bristol, Bristol, UK; Dr Christopher Emmett, County Durham and Darlington NHS Foundation Trust; Mr Ugo Grossi, Queen Mary University of London; Dr Richard Hooper, PCTU, Queen Mary University of London; Miss Emma Horrocks, University Hospital Southampton NHS Foundation Trust; Professor Charles Knowles, Queen Mary University of London; Mr Jon Lacy-Colson, Shrewsbury and Telford Hospital NHS Trust; Mr Ian Lindsey, Oxford University Hospitals NHS Trust; Professor James Mason, University of Warwick, Coventry; Mr Mark Mercer-Jones, Gateshead Health NHS Foundation Trust; Mr Andrew Miller, University Hospitals of Leicester NHS Trust; Mr David Pares, Hospital Germans Trias i Pujol, Barcelona, Spain; Miss Sophie Pilkington, University Hospital Southampton NHS Foundation Trust; Mr Neil Smart, Royal Devon & Exeter NHS Foundation Trust; Ms Natasha Stevens, PCTU, Queen Mary University of London; Professor Douglas Tincello, University Hospitals of Leicester NHS Trust; Miss Karen Telford, South Manchester NHS Foundation Trust; Mr Paul Vollebregt, Queen Mary University of London; Mr Andrew Williams, Guy's and Thomas' NHS Foundation Trust; Professor Yan Yiannakou, County Durham and Darlington NHS Foundation Trust.

Funding

This project was funded by the National Institute for Health Research (NIHR) Programme Grant for Applied Research (RP-PG-0612-20001). The views and opinions

expressed in this report are those of the authors and are not those of the PGfAR Programme, NIHR, the UK National Health Service, or the UK Department of Health.

Conflicts of interest

All authors have no conflict of interest in relation to the content of the manuscript. In the last 2 years, C.H. Knowles has received speaker fees and consultancy payments from Medtronic Inc., S. Brown has received speaker fees and consultancy payments from Allergan. M. Mercer-Jones has received fees as a preceptor for Medtronic Inc. and A.B. Williams has received speaker fees and consultancy payments from Medtronic Inc., and acts as a non-paid consultant for Cook Medical. M. Chapman has received funding to attend courses from Medtronic Inc.

References

- 1 Stewart WF, Liberman JN, Sandler RS *et al.* Epidemiology of constipation (EPOC) study in the United States: relation of clinical subtypes to sociodemographic features. *Am J Gastroenterol* 1999; **94**: 3530–40.
- 2 van den Berg MM, Benninga MA, Di Lorenzo C. Epidemiology of childhood constipation: a systematic review. *Am J Gastroenterol* 2006; **101**: 2401–9.
- 3 Soares 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**: 1582–91; quiz 1581, 1592.
- 4 Shafe AC, Lee S, Dalrymple JS, Whorwell PJ. The LUCK study: Laxative Usage in patients with GP-diagnosed Constipation in the UK, within the general population and in pregnancy. An epidemiological study using the General Practice Research Database (GPRD). *Therap Adv Gastroenterol* 2011; **4**: 343–63.
- 5 Knowles CH, Scott SM, Rayner C *et al.* Idiopathic slow-transit constipation: an almost exclusively female disorder. *Dis Colon Rectum* 2003; **46**: 1716–7.
- 6 Wald A, Scarpignato C, Mueller-Lissner S *et al.* A multinational survey of prevalence and patterns of laxative use among adults with self-defined constipation. *Aliment Pharmacol Ther* 2008; **28**: 917–30.
- 7 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. *Am J Gastroenterol* 2002; **97**: 1986–93.
- 8 Peery AF, Dellon ES, Lund J *et al.* Burden of gastrointestinal disease in the United States: 2012 update. *Gastroenterology* 2012; **143**: 1179–87.e1–3.
- 9 Poulton BTS. The nursing cost of constipation. *Primary Health Care* 1999; **9**: 17–22.
- 10 H.a.S.C.I.C. *Prescription Cost Analysis: England 2014*. Prescribing & Medicines Team, Health and Social Care Information Centre, 2015.

- 11 Palit S, Lunniss PJ, Scott SM. The physiology of human defecation. *Dig Dis Sci* 2012; **57**: 1445–64.
- 12 Ragg J, McDonald R, Hompes R, Jones OM, Cunningham C, Lindsey I. Isolated colonic inertia is not usually the cause of chronic constipation. *Colorectal Dis* 2011; **13**: 1299–302.
- 13 Bharucha AE, Wald A, Enck P, Rao S. Functional anorectal disorders. *Gastroenterology* 2006; **130**: 1510–8.
- 14 Woodward S, Norton C, Chiarelli P. Biofeedback for treatment of chronic idiopathic constipation in adults. *Cochrane Database Syst Rev* 2014; **3**: CD008486.
- 15 Liberati A, Altman DG, Tetzlaff J *et al.* The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *J Clin Epidemiol* 2009; **62**: e1–34.
- 16 Zorzela L, Loke YK, Ioannidis JP *et al.* PRISMA harms checklist: improving harms reporting in systematic reviews. *BMJ* 2016; **352**: i157.
- 17 Cook IJ, Talley NJ, Benninga MA, Rao SS, Scott SM. Chronic constipation: overview and challenges. *Neurogastroenterol Motil* 2009; **21**(Suppl 2): 1–8.
- 18 Maher C, Feiner B, Baessler K, Adams EJ, Hagen S, Glazner CM. Surgical management of pelvic organ prolapse in women. *Cochrane Database Syst Rev* 2010; **4**: CD004014.
- 19 Tou S, Brown SR, Nelson RL. Surgery for complete (full-thickness) rectal prolapse in adults. *Cochrane Database Syst Rev* 2015; **11**: CD001758.
- 20 Gladman MA, Scott SM, Lunniss PJ, Williams NS. Systematic review of surgical options for idiopathic megarectum and megacolon. *Ann Surg* 2005; **241**: 562–74.
- 21 Peng HY, Xu AZ. Colonic exclusion and combined therapy for refractory constipation. *World J Gastroenterol* 2006; **12**: 7864–8.
- 22 Knowles CH, Dinning PG, Pescatori M, Rintala R, Rosen H. Surgical management of constipation. *Neurogastroenterol Motil* 2009; **21**(Suppl 2): 62–71.
- 23 Moher D, Pham B, Lawson ML, Klassen TP. The inclusion of reports of randomised trials published in languages other than English in systematic reviews. *Health Technol Assess* 2003; **7**: 1–90.
- 24 Dekkers OM, Egger M, Altman DG, Vandembroucke JP. Distinguishing case series from cohort studies. *Ann Intern Med* 2012; **156**(Pt 1): 37–40.
- 25 Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions*. The Cochrane Collaboration, 2011. <http://handbook.cochrane.org>
- 26 Juillerat P, Peytremann-Bridevaux I, Vader JP *et al.* Appropriateness of colonoscopy in Europe (EPAGE II). Presentation of methodology, general results, and analysis of complications. *Endoscopy* 2009; **41**: 240–6.
- 27 Evans C, Ong E, Jones OM, Cunningham C, Lindsey I. Laparoscopic ventral rectopexy is effective for solitary rectal ulcer syndrome when associated with rectal prolapse. *Colorectal Dis* 2014; **16**: O112–6.

- 28 Finch KBS, Bernstein SJ, Aguilar MD *et al.* *The RAND/UCLA Appropriateness Method User's Manual*. Santa Monica, CA, USA: Rand Europe, 2001.
- 29 Brook RH, Chassin MR, Fink A, Solomon DH, Koseoff J, Park RE. A method for the detailed assessment of the appropriateness of medical technologies. *Int J Technol Assess Health Care* 1986; **2**: 53–63.

Appendix I: Search strategy

Population terms

“constipation”[All Fields] OR “obstructed defecation”[All Fields] OR “colonic inertia”[All Fields] OR “intussusception”[All Fields] OR “rectal prolapse”[All Fields] OR “outlet obstruction”[All Fields] OR “SRUS”[All Fields] OR “solitary rectal”[All Fields] OR “defecation disorder”[All Fields] OR “impaired defecation”[All Fields] OR “rectal emptying”[All Fields] OR “bowel dysfunction”[All Fields] OR “bowel function”[All Fields] OR “defecography”[All Fields] OR “defaecography”[All Fields] OR “defecographic”[All Fields] OR “evacuation difficulty”[All Fields] OR “evacuation disorder”[All Fields] OR (“Constipation”[Mesh Terms]) NOT (“child”[MeSH Terms]).

Intervention terms

“Delorme procedure”[All Fields] OR “delormes procedure”[All Fields] OR “delorme’s procedure”[All Fields] OR “sacral nerve stimulation”[All Fields] OR “sacral neuromodulation”[All Fields] OR “neurostimulation”[All Fields] OR “sacral nerve modulation”[All Fields] OR “STARR”[All Fields] OR “stapled transanal resection”[All Fields] OR “stapled transanal rectal resection” [All Fields] OR “trans-STARR”[All Fields] OR “Stapled trans-anal rectal resection”[All Fields] OR “rectopexy”[All Fields] OR “sacrocolpopexy”[All Fields] OR “sacropexy”[All Fields] OR “promontofixation”[All Fields] OR “colectomy”[All Fields] OR “proctocolectomy”[All Fields] OR “ileorectal”[All Fields] OR “cecorectal” [All Fields] OR “ileoproctostomy”[All Fields] OR “cecoproctostomy”[All Fields] OR “ileosigmoid”[All Fields] OR “rectocele repair”[All Fields] OR “posterior repair”[All Fields] OR “colporrhaphy”[All Fields] OR “rectovaginal septum reinforcement”[All Fields] OR “anterior rectal wall repair”[All Fields] OR “surgical repair of rectocele”[All Fields] OR (“transperineal mesh repair”[All Fields] OR “transperineal repair”[All Fields] OR “transvaginal repair”[All Fields] OR “transanal repair”[All Fields] OR “endorectal repair”[All Fields] OR “transrectal repair”[All Fields] OR “transanal longitudinal plication”[All Fields] OR (“Constipation/surgery”[Mesh Terms])

Report terms

(hasabstract[text]) AND (“0001/01/01”[PDat]: “2016/02/22”[PDat])

Appendix II: European reference group*

| Name | Institution |
|----------------------|--------------------|
| Donato Altomare | Bari, Italy |
| Lars Boenicke | Wuerzburg, Germany |
| Steen Buntzen | Tromsø, Norway |
| Ken Campbell | Dundee, UK |
| Peter Christensen | Aarhus, Denmark |
| Andre D'Hoore | Leuven, Belgium |
| Eloy Espin | Barcelona, Spain |
| David Jayne | Leeds, UK |
| Oliver Jones | Oxford, UK |
| Jens-Christian Knapp | Stavanger, Norway |
| Soren Laurberg | Aarhus, Denmark |
| Paul Lehur | Nantes, France |
| Klaus Matzel | Erlangen, Germany |
| Ronan O'Connell | Dublin, Ireland |
| Michel Prud'homme | Nimes, France |
| Carlo Ratto | Rome, Italy |
| Mario Trompetto | Turin, Italy |
| Caroline Vaizey | London, UK |

*Derived from European Society of Coloproctology.