
Publisher's PDF, also known as Version of record

License (if available):
CC BY

Link to published version (if available):
10.1111/codi.13021

Link to publication record in Explore Bristol Research

PDF-document

This is the final published version of the article (version of record). It first appeared online via Wiley at 10.1111/codi.13021. Please refer to any applicable terms of use of the publisher.

University of Bristol - Explore Bristol Research

General rights

This document is made available in accordance with publisher policies. Please cite only the published version using the reference above. Full terms of use are available: http://www.bristol.ac.uk/pure/user-guides/explore-bristol-research/ebr-terms/
Synthesis and summary of patient-reported outcome measures to inform the development of a core outcome set in colorectal surgery


*A Centre for Surgical Research, School of Social and Community Medicine, University of Bristol, Bristol, UK, †Severn School of Surgery, University Hospitals Bristol NHS Foundation Trust, Bristol, UK, ‡Division of Surgery Head and Neck, University Hospitals Bristol NHS Foundation Trust, Bristol, UK, §Colorectal Cancer Patient Representative, North Bristol NHS Trust, Bristol, UK, ¶Department of General Surgery, North Bristol NHS Trust, Bristol, UK, **Colorectal Surgery Unit, University Hospitals Bristol NHS Foundation Trust, Bristol, UK, ††Colorectal Consumer Liaison Group, National Cancer Research Institute, London, UK, †††Academic Department of Surgery, University of Birmingham, Birmingham, UK, §§Department of Surgery, St Mark’s Hospital and Academic Institute, Harrow, UK, §§§Academic Surgical Unit, St James’ University Hospital NHS Trust, Leeds, UK, ¶¶¶Centre for Ethics in Medicine, University of Bristol, Bristol, UK, ¶¶¶Colorectal Network Site Specific Group, Avon, Somerset and Wiltshire Cancer Services, Bristol, UK, ††††Centre for Statistics in Medicine and Oxford Clinical Trials Research Unit, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford, UK, §§§Department of Colorectal Surgery, Plymouth Hospitals NHS Trust, Plymouth, UK and ¶¶¶¶Clinical Trials Research Unit, University of Leeds, Leeds, UK

Received 16 March 2015; accepted 5 May 2015; Accepted Article online 8 June 2015

Abstract

Aim Patient-reported outcome (PRO) measures (PROMs) are standard measures in the assessment of colorectal cancer (CRC) treatment, but the range and complexity of available PROMs may be hindering the synthesis of evidence. This systematic review aimed to: (i) summarize PROMs in studies of CRC surgery and (ii) categorize PRO content to inform the future development of an agreed minimum ‘core’ outcome set to be measured in all trials.

Method All PROMs were identified from a systematic review of prospective CRC surgical studies. The type and frequency of PROMs in each study were summarized, and the number of items documented. All items were extracted and independently categorized by content by two researchers into ‘health domains’, and discrepancies were discussed with a patient and expert. Domain popularity and the distribution of items were summarized.

Results Fifty-eight different PROMs were identified from the 104 included studies. There were 23 generic, four cancer-specific, 11 disease-specific and 16 symptom-specific questionnaires, and three ad hoc measures. The most frequently used PROM was the EORTC QLQ-C30 (50 studies), and most PROMs (n = 40, 69%) were used in only one study. Detailed examination of the 50 available measures identified 917 items, which were categorized into 51 domains. The domains comprising the most items were ‘anxiety’ (n = 85, 9.2%), ‘fatigue’ (n = 67, 7.3%) and ‘physical function’ (n = 63, 6.9%). No domains were included in all PROMs.

Conclusion There is major heterogeneity of PRO measurement and a wide variation in content assessed in the PROMs available for CRC. A core outcome set will improve PRO outcome measurement and reporting in CRC trials.

Keywords Colorectal cancer, surgery, patient-reported outcomes, core outcome set, systematic review

Introduction

The measurement of patient-reported outcomes (PROs) has become standard in the assessment of colorectal cancer (CRC) treatments, and their use is recommended by
funding and regulatory agencies [1]. Many patient-reported outcome measures (PROMs) have therefore been developed for a variety of purposes [2]. Some are generic, and allow comparisons between patients with other conditions (e.g. SF-36, EQ-5D), others are designed for patients with cancer (e.g. EORTC QLQ-C30, FACT-G), and some are specific for CRC (e.g. EORTC-CR29, FACT-C). To add further complexity, each of these PROMs typically consists of a number of questions (items), which are often grouped together to represent similar concepts (scales). For example, two questions regarding activities of daily living and leisure activities in the EORTC QLQ-C30 measure are grouped into a single ‘role function’ scale. There are therefore a multitude of ways to measure PROs to evaluate treatment for CRC, and this creates problems that may influence the conduct and clinical impact of trials.

Trials may use different PROMs [3,4] making it impossible to synthesize data across trials or undertake meta-analyses. The multiplicity of results available from trials means that it is difficult to interpret findings in the context of clinical practice because of a lack of familiarity with the number of measures, scales and items [2]. For example, the scale ‘physical function’ exists in several different PROMs, but individual items in these scales vary considerably between questionnaires. This is confusing for clinicians, who may not be aware of the differences between PROMs, and it is likely to limit the meaningful use of the data in practice. Finally, the opportunities for measuring multiple outcomes may lead to selective reporting of significant findings. This can generate bias and influence clinical interpretation of trials [5].

A proposed solution to these issues are ‘core outcome sets’. Core outcomes are the minimum set of outcomes that patients and professionals agree should be measured in all trials of a certain condition [6]. They aim to facilitate comparisons between trials and aid meta-analysis by standardizing outcome measurement, including PROs. The use of core sets may also facilitate the clinical communication of data. Many core outcome sets have now been developed in different clinical areas, including rheumatology [7], paediatrics [8] and obstetrics [9], but not in CRC surgery. This systematic review aims to examine the measurement of PROs in CRC surgical studies, and use the data to inform the development of the core outcome set.

### Method

A systematic review of prospective CRC surgical studies measuring PROs was undertaken to: (i) summarize PRO measurement in CRC surgical studies, and (ii) examine each PROM in detail and categorize analogous

<table>
<thead>
<tr>
<th>Table 1</th>
<th>OvidSP version of Medline search strategy.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Search criteria</td>
<td>Search terms</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>1. exp Colonic Neoplasms/ 2. exp Rectal Neoplasms/ 3. ((colorect$ or colon or colonic or rect$) adj3 (cancer$ or tumor$ or neoplasm$ or carcinoma$ or malignan$)).tw. 4. or/1-3</td>
</tr>
<tr>
<td>Surgery</td>
<td>1. exp Specialties, Surgical/ 2. surg$.tw. 3. operat$.tw. 4. intervention$.tw. 5. procedur$.tw. 6. resect$.tw. 7. or/1-6</td>
</tr>
<tr>
<td>Randomized controlled trials/ prospective studies</td>
<td>1. randomized controlled trial.pt. 2. controlled clinical trial.pt. 3. randomized controlled trials.sh. 4. random allocation.sh. 5. double blind method.sh. 6. single-blind method.sh. 7. or/1-6 8. exp animals/not human/ 9. 7 not 8 10. clinical trial.pt. 11. exp clinical trials/ 12. (clin$ adj25 trial$).ti,ab. 13. ((singl$ or doubl$ or trebl$ or tripl$) adj25 (blind$ or mask$)).ti,ab. 14. placebos.sh. 15. placebo$.ti,ab. 16. random$.ti,ab. 17. research design.sh. 18. or/10-17 19. 18 not 8 20. 19 not 9 21. comparative study.sh. 22. exp evaluation studies/ 23. follow up studies.sh. 24. prospective studies.sh. 25. (control$ or prospective$ or volunteer$).ti,ab. 26. or/21-25 27. 26 not 8 28. 27 not (9 or 20) 29. 9 or 20 or 28</td>
</tr>
</tbody>
</table>

This systematic review adhered to a predefined protocol (available on request from the authors). Validated terms concepts into domains to inform the future development of a core outcome set.

### Systematic search and data extraction

This systematic review adhered to a predefined protocol (available on request from the authors). Validated terms
relating to ‘surgery’, ‘colorectal cancer’ and ‘prospective studies’ (Table 1) were used to search the OVID SP versions of MEDLINE, EMBASE and the Cochrane Central Register of Controlled Trials. A validated filter for ‘prospective studies’ was used because PRO data are typically collected prospectively. The search was limited to studies conducted in humans aged 18 years and over, reported in the English language between January 2009 and December 2010. Previous reviews have considered PROs of CRC surgery in terms of elderly patients [10], methodological challenges in measuring PROs in CRC [11], laparoscopic surgery [12], long-term survivors [13], rectal cancer [3] and CRC before 2009 [14]. The studies identified in these reviews were included. All citations were collated with REFERENCE MANAGER 12 (Thomson Reuters, New York city, New York, USA) and the duplicates removed.

Titles and abstracts of identified publications were screened by one researcher. If there was uncertainty about the eligibility of a publication the full paper was also accessed. Articles were included if they were original research papers reporting PROs of CRC surgery (curative or palliative), with or without neoadjuvant or adjuvant therapies, or systematic reviews of such publications. PROs were defined as end-points provided by patients themselves and not interpreted by observers. Studies of nonbiomedical interventions (e.g. alternative medicine), palliative treatments that did not include a surgical component (e.g. palliative chemotherapy), screening studies, treatment of colorectal metastases and molecular and genetic prognostic studies were all excluded. Studies of more than one cancer site or of mixed benign and malignant disease were included provided the data for CRC patients was presented separately from that of the other diseases.

Data extraction included: participant demographics (number, age and gender); treatment received (surgery, neoadjuvant radiotherapy/chemoradiation and adjuvant chemotherapy); treatment intent (curative or palliative); the study design (randomized trial, case–control study, cohort study, cross-sectional study, prospective case series or other design); the PRO questionnaire used; and the individual items included in each questionnaire. When the individual questionnaires were not available in publications, internet searches and direct contact with authors were used to obtain the information. All data were entered into a Microsoft Access (Microsoft, Redmond, Washington, USA) database to facilitate data management and analyses. The data extraction was checked by a second reviewer (ROF) for a sample of included articles (n = 25) and any disagreements were discussed and resolved with the senior author (JMB).
Summary of PRO measurement in CRC surgery

The number of publications reporting each PROM was tabulated and descriptive statistics used to summarize PRO measurement. The popularity of PROMs was assessed by comparing their frequency of use in studies. A summary of each PROM is provided in terms of the numbers of items, scales and whether a total score was used. The distribution of items among PROMs was examined by calculating the median number and range of items per PROM. Questionnaires were categorized as: (i) generic (for use in all patients), (ii) cancer specific (for use in all cancer patients), (iii) CRC specific (for use in CRC patients), (iv) symptom specific (to assess a single symptom, e.g. pain), or (v) ad hoc.

Examination of PROs and domain categorization

Individual items from all questionnaires were extracted and formed into a long-list before categorization into health domains by two researchers (RNW and JR). Both were kept masked as to which PRO questionnaire the items were derived from. Two patient representatives (JEJ and GS) and one consultant colorectal surgeon (AMP) subsequently checked this process. Discrepancies were discussed and resolved with the senior author (JMB).

Table 2 Summary of included articles.

<table>
<thead>
<tr>
<th></th>
<th>All studies (n = 102)</th>
<th>Randomized trials (n = 25)</th>
<th>Nonrandomized studies (n = 77)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of participants</td>
<td>66 386</td>
<td>7172</td>
<td>59 214</td>
</tr>
<tr>
<td>Age range of participants (years)</td>
<td>18–99</td>
<td>29–89</td>
<td>18–99</td>
</tr>
<tr>
<td>Number of participating centres (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Single</td>
<td>58 (57)</td>
<td>13 (52)</td>
<td>45 (58)</td>
</tr>
<tr>
<td>Multiple</td>
<td>44 (43)</td>
<td>12 (48)</td>
<td>32 (42)</td>
</tr>
<tr>
<td>IRB or ethical approval reported (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colon</td>
<td>10 (9)</td>
<td>2 (8)</td>
<td>7 (9)</td>
</tr>
<tr>
<td>Rectum</td>
<td>54 (53)</td>
<td>11 (44)</td>
<td>44 (55)</td>
</tr>
<tr>
<td>Mixed colon and rectum</td>
<td>38 (38)</td>
<td>12 (48)</td>
<td>26 (34)</td>
</tr>
<tr>
<td>Surgical approach (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Laparoscopy</td>
<td>1 (1)</td>
<td>0</td>
<td>1 (1)</td>
</tr>
<tr>
<td>Hand-assisted laparoscopy</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Open</td>
<td>5 (5)</td>
<td>1 (4)</td>
<td>4 (5)</td>
</tr>
<tr>
<td>Mixed</td>
<td>13 (12)</td>
<td>9 (36)</td>
<td>5 (6)</td>
</tr>
<tr>
<td>Not reported or incomplete information reported</td>
<td>83 (81)</td>
<td>15 (60)</td>
<td>67 (87)</td>
</tr>
<tr>
<td>Neoadjuvant treatment* (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Radiotherapy alone</td>
<td>20 (20)</td>
<td>9 (36)</td>
<td>11 (14)</td>
</tr>
<tr>
<td>Chemotherapy alone</td>
<td>0</td>
<td>0</td>
<td>0v (0.0)</td>
</tr>
<tr>
<td>Chemoradiotherapy</td>
<td>22 (22)</td>
<td>3 (12)</td>
<td>18 (24)</td>
</tr>
<tr>
<td>None</td>
<td>7 (7)</td>
<td>1 (4)</td>
<td>6 (8)</td>
</tr>
<tr>
<td>Not reported or incomplete information reported</td>
<td>58 (51)</td>
<td>12 (48)</td>
<td>42 (54)</td>
</tr>
<tr>
<td>Adjuvant treatment* (%)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chemotherapy or chemoradiotherapy</td>
<td>56 (55)</td>
<td>16 (64)</td>
<td>40 (52)</td>
</tr>
<tr>
<td>None</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Not reported or incomplete information reported</td>
<td>46 (45)</td>
<td>9 (36)</td>
<td>37 (48)</td>
</tr>
<tr>
<td>Number of PROMs reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>43</td>
<td>10</td>
<td>33</td>
</tr>
<tr>
<td>2</td>
<td>47</td>
<td>12</td>
<td>35</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>0</td>
<td>2</td>
</tr>
</tbody>
</table>

IRB, Institutional Review Board.
*Some studies included patients with or without neoadjuvant therapy, some patients underwent different neoadjuvant or adjuvant treatment within the same study.
Table 3 Summary of identified patient-reported outcome measures (questionnaires) (n = 58).

<table>
<thead>
<tr>
<th>Name of generic questionnaire (n = 23)</th>
<th>Number of items</th>
<th>Number of scales</th>
<th>Overall score</th>
<th>Frequency (n = 184)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short Form-36</td>
<td>36</td>
<td>8</td>
<td>No</td>
<td>21</td>
</tr>
<tr>
<td>EuroQol-5D</td>
<td>6</td>
<td>6</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Rotterdam Symptom Checklist</td>
<td>35</td>
<td>4</td>
<td>No</td>
<td>3</td>
</tr>
<tr>
<td>Gastrointestinal Quality of Life Index</td>
<td>36</td>
<td>0</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Functional Difficulty Index</td>
<td>15</td>
<td>0</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Illness Impact Scale</td>
<td>9</td>
<td>0</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Visual Analogue Scale (overall health)</td>
<td>1</td>
<td>0</td>
<td>Yes</td>
<td>2</td>
</tr>
<tr>
<td>Self-rated health*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Freiburger Illness Coping Strategies questionnaire*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Brief Symptom Inventory-18</td>
<td>18</td>
<td>3</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Constructed Meaning Scale</td>
<td>8</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Surgical Recovery Score</td>
<td>31</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Nottingham Health Profile</td>
<td>45</td>
<td>6</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Duke Generic Instrument</td>
<td>17</td>
<td>11</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Instrumental Activities of Daily Living</td>
<td>7</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Profile of Moods States</td>
<td>65</td>
<td>6</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Health and Activities Limitation Index</td>
<td>8</td>
<td>2</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Health Utility Index</td>
<td>7</td>
<td>7</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Spitzer Quality of Life Index</td>
<td>5</td>
<td>5</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Global Quality of Life*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Multidimensional Functional Assessment Questionnaire*</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Symptom Experience Scale</td>
<td>24</td>
<td>6</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Ad hoc satisfaction questionnaire†</td>
<td>6</td>
<td>6</td>
<td>Yes</td>
<td>1</td>
</tr>
</tbody>
</table>

Name of cancer-specific questionnaire (n = 4)

| EORTC QLQ-C30                          | 30              | 15              | No            | 50                  |
| Cancer-related Health Worries Scale    | 4               | 0               | Yes           | 2                   |
| Quality of Life – Cancer Survivors    | 41              | 4               | No            | 1                   |
| Cancer Problems in Living Scale       | 31              | 0               | Yes           | 1                   |

Name of disease-specific questionnaire (n = 11)

| EORTC QLQ-CR38                        | 38              | 9               | No            | 33                  |
| Functional Assessment of Cancer Therapy – Colorectal | 37          | 5               | Yes           | 5                   |
| Modified City of Hope Quality of Life – Ostomy | 41          | 6               | Yes           | 2                   |
| EORTC QLQ-CR29                        | 34              | 4               | No            | 1                   |
| University of Padova Bowel Function Questionnaire | 8              | 0               | Yes           | 1                   |
| Bowel Function Questionnaire           | 8               | 0               | Yes           | 1                   |
| Bowel Problems Scale                   | 7               | 7               | No            | 1                   |
| Late Effects Normal Tissue – subjective, objective, management, analytic scale* | – | – | – | 1 | |
| Quality of Life Index for Colostomy Patients | 23          | 3               | No            | 1                   |
| Colorectal Cancer Quality of Life     | 62              | 4               | Yes           | 1                   |
| COloREctal Functional Outcome Questionnaire | 26         | 5               | Yes           | 1                   |

Name of symptom-specific questionnaire (n = 17)

| International Index of Erectile Function | 15          | 5               | Yes           | 4                   |
| Faecal Incontinence Quality of Life Questionnaire | 29          | 4               | No            | 3                   |
| Wexner Incontinence Scale               | 5               | 0               | Yes           | 3                   |
| Visual Analogue Scale (pain)            | 1               | 0               | Yes           | 3                   |
| Center for Epidemiologic Studies – Depression | 20          | 6               | Yes           | 3                   |
| Hospital Anxiety and Depression Scale   | 14              | 0               | Yes           | 2                   |
| Holschneider Questionnaire              | 8               | 0               | Yes           | 1                   |
| Internation Index of Erectile Function-5 | 5              | 5               | Yes           | 1                   |
| Body Image Questionnaire                | 10              | 2               | No            | 1                   |
Categorization was summarized using descriptive statistics to explore the distribution of items and PROMs between domains. The number of items included in each domain was counted, as were the number of PROMs from which they were sourced. The contribution of each source PROM was demonstrated by calculating the median number and range of items included from the measures.

**Results**

A total of 5644 titles and abstracts were identified, of which 2127 were duplicates. The remaining 3517 were screened and 29 original research articles included. In addition to this, six systematic reviews of PROs in CRC surgery identified a further 72 original research articles ([Fig. 1](#)). In total, 102 original publications including 25 randomized controlled trials (25%) and 77 nonrandomized studies (75%) reporting the outcome for 66 386 patients with CRC [15–116] were included. The studies are summarized in [Table 2](#).

<table>
<thead>
<tr>
<th>Name of questionnaire (n = 3)</th>
<th>Number of items</th>
<th>Number of scales</th>
<th>Overall score</th>
<th>Frequency (n = 184)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Body Image Scale</td>
<td>10</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Faecal Incontinence Scoring System</td>
<td>5</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Patient Assessment of Constipation Symptom Scale</td>
<td>12</td>
<td>3</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Present Pain Intensity Index</td>
<td>1</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Satisfaction with Sexual Function</td>
<td>1</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Visual Analogue Scale (wound satisfaction)</td>
<td>1</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Symptom Distress Scale</td>
<td>15</td>
<td>0</td>
<td>Yes</td>
<td>1</td>
</tr>
<tr>
<td>Multidimensional Fatigue Inventory-20</td>
<td>20</td>
<td>5</td>
<td>No</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Name of questionnaire (n = 3)</th>
<th>Number of items</th>
<th>Number of scales</th>
<th>Overall score</th>
<th>Frequency (n = 184)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ad hoc QOL questionnaire A*†</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Ad hoc QOL questionnaire B*†</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
<tr>
<td>Ad hoc QOL questionnaire C*†</td>
<td>–</td>
<td>–</td>
<td>–</td>
<td>1</td>
</tr>
</tbody>
</table>

EORTC, European Organisation for Research and Treatment of Cancer; QOL, quality of life.
*Questionnaire not available.
†PROM not validated.

Most questionnaires were reported only once (n = 40, 69%). The most frequently reported were the European Organisation for the Research and Treatment of Cancer (EORTC) QLQ-C30 (50 studies, 48%), the EORTC QLQ-CR38 (33 studies, 32%) and the Medical Outcome Study Short Form-36 (21 studies, 21%). The median number of items per PROM was 14, ranging from one [five PROs: Visual Analogue Scale (overall, pain and wound satisfaction), Satisfaction with Sexual Function, and Present Pain Intensity Index] to 65 (the Profile of Mood States). Some 159 scales were evident, and most PROMs (n = 48, 83%) included a total score.

**Examination of PROs and domain categorization**

Fifty (86.2%) full questionnaires were available. Reasons for unavailability were inability to obtain the questionnaires from authors or web searches (n = 6) or lack of an English language translation (n = 2). The 50 questionnaires comprised some 917 individual items, and were categorized into 51 domains as described above ([Table 4](#)). The full categorization is presented in [Table S1](#). The domains comprising the most items were ‘anxiety’ (n = 85, 9.2%), ‘fatigue’ (n = 67, 7.2%) and ‘physical function’ (n = 63, 6.8%). The disease-specific domains comprising most items were ‘faecal incontinence’ (n = 53, 5.7%) and ‘stoma problems’ (n = 52, 5.6%). Most domains (n = 27, 53%) contained 10 or more items.

There was little evidence of consistency between PROMs. No domains were measured in all the PROMs.
Table 4  Summary of domain categorization including number of items per domain, numbers of PROMs and median items per PROM.

<table>
<thead>
<tr>
<th>PRO domain (n = 51)</th>
<th>Number of items (n = 917) (%)</th>
<th>Number of PROMs (n = 50) (%)</th>
<th>Median items per source PROM (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Psychological domains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anxiety</td>
<td>85 (9.2)</td>
<td>22 (44)</td>
<td>2.5 (1–12)</td>
</tr>
<tr>
<td>Fatigue</td>
<td>67 (7.2)</td>
<td>21 (42)</td>
<td>1.0 (1–23)</td>
</tr>
<tr>
<td>Depression</td>
<td>47 (5.1)</td>
<td>16 (32)</td>
<td>1.5 (1–12)</td>
</tr>
<tr>
<td>Body image</td>
<td>37 (4.0)</td>
<td>13 (26)</td>
<td>1.0 (1–10)</td>
</tr>
<tr>
<td>Frustration/irritability</td>
<td>15 (1.6)</td>
<td>7 (14)</td>
<td>1.0 (1–9)</td>
</tr>
<tr>
<td>Outlook on life</td>
<td>13 (1.4)</td>
<td>5 (10)</td>
<td>2.0 (1–6)</td>
</tr>
<tr>
<td>Self-esteem</td>
<td>11 (1.2)</td>
<td>6 (12)</td>
<td>2.0 (1–3)</td>
</tr>
<tr>
<td>Coping</td>
<td>10 (1.1)</td>
<td>6 (12)</td>
<td>1.0 (1–3)</td>
</tr>
<tr>
<td>Spiritual</td>
<td>7 (0.7)</td>
<td>2 (4)</td>
<td>3.5 (3–4)</td>
</tr>
<tr>
<td>Regret</td>
<td>5 (0.5)</td>
<td>2 (4)</td>
<td>2.5 (1–4)</td>
</tr>
<tr>
<td>Control</td>
<td>3 (0.3)</td>
<td>3 (6)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td><strong>Functional domains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Physical function</td>
<td>63 (6.8)</td>
<td>19 (38)</td>
<td>1.0 (1–9)</td>
</tr>
<tr>
<td>Role function</td>
<td>51 (5.5)</td>
<td>20 (40)</td>
<td>2.0 (1–7)</td>
</tr>
<tr>
<td>Social function</td>
<td>50 (5.4)</td>
<td>22 (44)</td>
<td>2.0 (1–8)</td>
</tr>
<tr>
<td>Sexual function</td>
<td>44 (4.7)</td>
<td>13 (26)</td>
<td>1.0 (1–15)</td>
</tr>
<tr>
<td>Cognitive function</td>
<td>30 (3.2)</td>
<td>14 (28)</td>
<td>1.0 (1–7)</td>
</tr>
<tr>
<td><strong>Symptom domains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Faecal incontinence</td>
<td>53 (5.7)</td>
<td>12 (24)</td>
<td>2.0 (1–27)</td>
</tr>
<tr>
<td>Stoma problems</td>
<td>52 (5.6)</td>
<td>5 (10)</td>
<td>7.0 (7–21)</td>
</tr>
<tr>
<td>Pain</td>
<td>50 (5.4)</td>
<td>18 (36)</td>
<td>1.5 (1–8)</td>
</tr>
<tr>
<td>Insomnia</td>
<td>18 (1.9)</td>
<td>13 (26)</td>
<td>1.0 (1–4)</td>
</tr>
<tr>
<td>Appetite/eating problems</td>
<td>17 (1.8)</td>
<td>10 (20)</td>
<td>1.5 (1–3)</td>
</tr>
<tr>
<td>Faecal frequency</td>
<td>14 (1.5)</td>
<td>8 (16)</td>
<td>2.0 (1–3)</td>
</tr>
<tr>
<td>Nausea/vomiting</td>
<td>12 (1.3)</td>
<td>8 (16)</td>
<td>1.0 (1–3)</td>
</tr>
<tr>
<td>Faecal Urgency</td>
<td>11 (1.2)</td>
<td>8 (16)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Flatulence or gas</td>
<td>11 (1.2)</td>
<td>7 (14)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Treatment problems</td>
<td>11 (1.2)</td>
<td>7 (14)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Rectal blood or mucus</td>
<td>10 (1.1)</td>
<td>8 (16)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Bloating</td>
<td>7 (0.7)</td>
<td>6 (12)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Diarrhoea</td>
<td>7 (0.7)</td>
<td>7 (14)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Tenesmus</td>
<td>7 (0.7)</td>
<td>4 (8)</td>
<td>2.0 (1–2)</td>
</tr>
<tr>
<td>Constipation</td>
<td>6 (0.6)</td>
<td>5 (10)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Shortness of breath</td>
<td>5 (0.5)</td>
<td>5 (10)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Urinary frequency</td>
<td>5 (0.5)</td>
<td>3 (6)</td>
<td>2.0 (1–2)</td>
</tr>
<tr>
<td>Faint or dizzy</td>
<td>4 (0.4)</td>
<td>4 (8)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Hair problems</td>
<td>4 (0.4)</td>
<td>4 (8)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Discrimination</td>
<td>3 (0.3)</td>
<td>3 (6)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Dry mouth</td>
<td>3 (0.3)</td>
<td>3 (6)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Menstruation</td>
<td>3 (0.3)</td>
<td>3 (6)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Taste</td>
<td>3 (0.3)</td>
<td>3 (6)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Duration of bowel movement</td>
<td>2 (0.2)</td>
<td>2 (4)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>2 (0.2)</td>
<td>2 (4)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Dysphagia</td>
<td>2 (0.2)</td>
<td>1 (2)</td>
<td>2.0 (2)</td>
</tr>
<tr>
<td>Dysuria</td>
<td>1 (0.1)</td>
<td>1 (2)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td>Urinary incontinence</td>
<td>1 (0.1)</td>
<td>1 (2)</td>
<td>1.0 (1)</td>
</tr>
<tr>
<td><strong>Global domains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Global quality of life</td>
<td>12 (1.3)</td>
<td>9 (18)</td>
<td>1.0 (1–2)</td>
</tr>
<tr>
<td>Self-care</td>
<td>10 (1.1)</td>
<td>10 (20)</td>
<td>1.0 (1)</td>
</tr>
</tbody>
</table>
The two domains that were best represented were ‘anxiety’ and ‘social function’, each measured by 22 (44%) PROMs. Otherwise, most domains \( (n = 39, 76\%) \) were measured by less than a quarter of PROMs, highlighting further heterogeneity. There were two domains with a high median number of items included per PROM: ‘stoma problems’, which contained 52 items from only five PROMs (median seven items per PROM) and ‘satisfaction with care’, which featured six items from just one PROM. This may reflect specialization of PROMs, with some measures focusing on very specific concepts.

**Discussion**

This systematic review aimed to summarize PRO measurement in current CRC surgical studies and categorize PRO items into analogous concepts to inform the development of a core outcome set. There was evidence of significant heterogeneity of PRO measurement in the included studies. Fifty-eight different PROMs were used to assess patient experience of colorectal surgery. Most \( (n = 40, 69.0\%) \) were only ever used once, and even the most common (EORTC QLQ-C30) was measured in less than half of the studies. PROMs also varied greatly in terms of their content, with some as simple as a single item while others included up to 65. Most (52%) PROMs were not designed to be specific to CRC surgery or symptoms thereof, and although this may bring benefits in terms of comparison between diseases they may not be sensitive enough to issues that are of specific importance to patients with CRC. Over 900 individual questionnaire items were evident from 50 PROMs, and through a rigorous process, these were categorized into 51 ‘health domains’. This demonstrated a further lack of consistency, with no domains being measured in all the PROMs, and most health domains only being measured by less than a quarter of PROMs. All of this highlights potentially major questions for evidence synthesis and clinical interpretation of results in studies of CRC surgery, and demonstrates the need for a standardized core outcome set.

Other studies have highlighted the problem with outcome heterogeneity for clinical and PRO data. A recently published systematic review identified 194 studies of CRC surgery that measured 766 different clinical outcomes, with no single outcome reported in all [117]. Even considering a seemingly simple outcome such as mortality, there were over 84 different ways in which this was defined and measured. The same problem has been highlighted in studies of oesophageal cancer surgery [118], where a review of 122 articles reported 210 unique complications and 10 different measures of operative mortality, and breast reconstruction following mastectomy for cancer [119], which identified 134 studies reporting 950 unique complications. Problems with the multiplicity of PRO measures have also been described previously in oesophageal surgery [120], but there is no evidence of this issue in trials of CRC surgery.

This study is the largest systematic review of PROs in CRC and was conducted with rigorous methodology, but there are some limitations. The review covers published CRC studies in English up until 2010. A more exhaustive search over a more recent period of time, or inclusion of unpublished data or non-English publications may have yielded further PROs, but all the most commonly used PROs were captured by these inclusion criteria and extending the review would have probably only identified additional rare PROs. The categorization process could be criticized as arbitrary, but efforts were made to objectify the process. First, two researchers categorized the questionnaire items independently, each blinded to the other. Second, categorization was checked for face validity by a patient representative. Finally, there has been full disclosure of the categorization in this article to allow scientific scrutiny of the process.

Having identified all the potential patient reported health domains measured in CRC surgical studies, the next phase of this research is to gain a consensus on which outcomes it is essential to measure in all trials. Recommended methods to achieve this have been defined by the international Core Outcome Measurement in Effectiveness Trials (COMET) group [6]. Domains will be combined with clinical outcomes generated from a previous systematic review [117] to create

**Table 4** (Continued).

<table>
<thead>
<tr>
<th>PRO domain ( (n = 51) )</th>
<th>Number of items ( (n = 917 %) )</th>
<th>Number of PROMs ( (n = 50 %) )</th>
<th>Median items per source PROM (range)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Financial</td>
<td>8 (0.9)</td>
<td>5 (10)</td>
<td>1.0 (1–4)</td>
</tr>
<tr>
<td>Satisfaction with care</td>
<td>6 (0.6)</td>
<td>1 (2)</td>
<td>6.0 (6)</td>
</tr>
<tr>
<td>Information needs</td>
<td>1 (0.1)</td>
<td>1 (2)</td>
<td>1.0 (1)</td>
</tr>
</tbody>
</table>

\( \text{ª2015 The Authors. Colorectal Disease published by John Wiley & Sons Ltd.} \)
an exhaustive long-list of all potential outcomes. Key stakeholders, including patients and professionals, will then consider the importance of these outcomes and undertake a prioritization exercise called the Delphi process. This will allow the outcomes of lesser importance to be discarded from the core set. Finally, when the number of outcomes has been reduced, face-to-face meetings will be conducted to allow for debate about their relative merits before the final core set is agreed.

In conclusion, this systematic review of CRC surgery demonstrated significant heterogeneity of PRO measurement that may hinder comparisons between studies, limit meta-analysis and allow outcome reporting bias. A long-list of patient reported ‘health domains’ was generated using robust methodology to inform the development of a core outcome set.

References

Synthesis and summary of PROMs


Synthesis and summary of PROMs


**Supporting Information**

Additional Supporting Information may be found in the online version of this article:

**Table S1.** Full categorization of patient reported outcome items.