---

**Monograph:**
Ara, R., Duarte, A., Harnan, S. et al. (4 more authors) (2015) . Supporting the routine collection of patient reported outcome measures in the National Clinical Audits for assessing cost-effectiveness. Work Package 1. What patient reported outcome measures should be used in the 2013/14 National Clinical Audit Programme. Appendix F - Bowel Cancer. Report. Policy Research Unit in Economic Evaluation of Health and Care Interventions (EEPRU) (RR0032). Policy Research Unit in Economic Evaluation of Health and Care Interventions

---

EEPRU

Economic Evaluation of Health and Social Care Interventions Policy Research Unit

# RESEARCH REPORT

## Supporting the routine collection of patient reported outcome measures in the National Clinical Audits for assessing cost-effectiveness

Work Package 1

What patient reported outcome measures should be used in the 13 health conditions specified in the 2013/14 National Clinical Audit programme?

APPENDIX F, BOWEL CANCER

Authors: Roberta Ara, Ana Duarte, Sue Harnan, Jo Leaviss, Steve Palmer, Mark Sculpher, John Brazier

Correspondence to: Katherine Stevens, HEDS, ScHARR, University of Sheffield, Regent Court, 30 Regent Street, Sheffield, S1 4DA. Email: r.m.ara@sheffield.ac.uk

RR0032

July 2015

The University Of Sheffield.

UNIVERSITY of York

The Policy Research Unit in Economic Evaluation of Health and Care interventions is funded by the Department of Health Policy Research Programme. It is a collaboration between researchers from the University of Sheffield and the University of York.

The Department of Health's Policy Research Unit in Economic Evaluation of Health and Care Interventions is a 7 year programme of work that started in January 2011. The unit is led by Professor John Brazier (Director, University of Sheffield) and Professor Mark Sculpher (Deputy Director, University of York) with the aim of assisting policy makers in the Department of Health to improve the allocation of resources in health and social care.

INDEX



TABLES



FIGURES

| Acronym | Definition |
|---|---|
| AE | Adverse events |
| AMSTAR | Assessing the quality of systematic reviews |
| APR | Abdominoperineal resection |
| ASA | American Society of Anesthesiologists classification, |
| BMI | Body mass index (kg/$m^2$) |
| CG | Clinical guideline |
| CPA | Colo-anal J-pouch anastomosis |
| DH | Department of Health |
| ECOG | Eastern Cooperative Oncology Group |
| EORTC-QLQ | European organisation for research and treatment of cancer core quality of life questionnaire |
| EQ-5D | EuroQol 5 dimensions |
| FACT-C | Functional assessment of cancer therapy - colorectal |
| FACT-EW | Functional assessment of cancer therapy – emotional well-being |
| FCSI | FACT colorectal symptom index |
| FIQL | Faecal incontinence quality of life |
| FISI | Faecal incontinence severity index |
| FR | Future research |
| FU | Fluorouracil |
| HADS | Hospital anxiety and depression scale |
| HRQoL | Health related quality of life |
| HS | Health states |
| HTA | Health technology assessment |
| HUI | Health Utility Index |
| HUI2 | Health Utility Index mark 2 |
| HUI3 | Health Utility Index mark 3 |
| LCR | Laparoscopic colon resection |
| LRA | Transanally double stapled low colorectal anastomosis |
| LV | Leucovorin |
| MDT | Multi-disciplinary team |
| MTA | Multiple technology assessment |
| NCA | National Clinical Audit |
| NHS | National Health Service |
| NICE | National Institute for Health and Care Excellence |
| OS | Overall survival |
| PANAS | Positive and negative affect scale |
| PANSS | Positive and negative syndrome scale |
| PFS | Progression free survival |
| PR | Potential recommendations |
| PROM(s) | patient reported outcome measure(s) |
| R&D | Research and development |
| RCT | Randomised controlled trial |
| RSCL | Rotterdam symptom checklist |
| SF-6D | Short form 36 |
| SG | Standard gamble |
| STA | Single technology assessment |
| TA | Technology Appraisal |

| | |
|---|---|
| TAG | Technology Assessment group |
| TEM | Transanal endoscopic microsurgery |
| TME | Total mesorectal excision |
| TNM | Tumour, node, metastases |
| TTO | Time trade off |
| UK | United Kingdom |
| VAS | Visual analogue scale |
| WP | Work package |

## 1. BACKGROUND

EEPRU was approached by Jason Cox (R&D Division) to prepare a programme of research to support the appropriateness of, and use of, patient reported outcome measures (PROMs) collected for the National Clinical Audit (NCA). The EEPRU programme was informed by a Research and Development (R&D) template prepared by Simon Bennett, Steve Fairman and Keith Willett at NHS England.

The purpose of introducing PROMs into the NCA programme is to be able to 1) compare performance between providers and commissioners in the National Health Service (NHS), 2) compare the cost-effectiveness of alternative providers in delivering the specific services (i.e. linking outcomes and resource use), and 3) assess the cost-effectiveness of alternative interventions and other changes in the NHS. The intention is to introduce PROMs across a range of conditions over the next 3 years commencing with 13 conditions in the 2014/15 NCA programme.

The agreed research programme consists of 3 concurrent work packages (WP) as described in the document submitted to the DH (8th November 2013). The current document provides details on the objectives, methodology and results for Work Package 1 (WP1): to determine what PROMS should be used in the 13 health conditions specified in the 2014/15 NCA programme.

## 2. OVERVIEW

WP1 is split into three separate components consisting of:

- WP1.1 To examine whether the EQ-5D is appropriate in the 13 health conditions specified in the 2013/14 NCA programme.
- WP1.2 To identify what measure could be used when the EQ-5D is not appropriate in the 13 health conditions, taking into account that the proposed measure would be used to generate preference-based utility measures (either directly through existing preference-based weights, or indirectly through existing mapping functions suitable for the proposed measure).
- WP1.3 To identify the evidence required to address questions of cost-effectiveness using the NCA data.

Each component consists of a series of reviews of the literature.

This Appendix provides the detailed results for the condition bowel cancer and should be read in conjunction with both the main report and the methods/search strategy appendices.

## 3. METHOD

The full detailed methodology used is provided in Appendix A, including the search strategy, selection criteria for studies included, and data extraction etc. In summary, a review of the literature was undertaken to assess the appropriateness of the EQ-5D in terms of classic psychometric criteria (WP1.1); where the EQ-5D was not considered appropriate, additional searches were undertaken to identify alternative measures (WP1.2); and finally, existing health technology appriasials were reviewed and data requirements were compared with variables currently collected in the IBD audit (WP1.3).

### 3.1 Psychometric properties (WP1.1)

Assessments reported in the included studies were categorised according to the following definitions:

*Acceptability*

Data relating to how acceptable the measure was to the person completing it, expressed as the proportion of completed surveys, or the proportion of missing data.

*Reliability*

There are two main definitions for reliability, a) the degree to which a measure reproduces the same results in an unchanged population and b) the degree to which a measure reproduces the same results when completed by different assessors (e.g. patient and proxy report). In both cases, reliability can be assessed by re-testing, and calculating the correlations or difference between tests. In case a) the comparison may be between the same populations separated by time, where no change in health state was observed (as compared to using an alternative condition specific or generic measure). In case b) the measure may be completed by multiple people (proxies) on the patient's behalf and their responses compared with those of the patient. Where the outcome measure is specifically designed for self-report by patients, this test of reliability may be expected to produce less agreement.

*Construct validity*

This is an assessment of how well an instrument measures what it intends to measure. Two main definitions are used in this review.

*a) Known group validity*, where estimates for groups that are known to differ in a concept of interest are compared either qualitatively or statistically. The known groups may be defined using other measures, according to clinical categorisation.

*b) Convergent validity* assesses the extent to which a measure correlates with other measures of the same or similar concepts. Correlation coefficients were considered low if <0.3, moderate if between 0.3 and 0.5, and strong when >0.5.

*Responsiveness*

*a) Change over time.* This is an assessment of whether measurements using the instrument can detect a change over time, where a change is expected. This may be before and after an intervention, or through progression of a disease. Evidence was considered to be good where a t-test was significant, though weaker evidence to support responsiveness was considered where there was a change in the expected direction, but was not statistically significant or not tested. Effect size and standardised response mean were also acceptable assessments of responsiveness.

*b) Ceiling and floor effects* were also considered to be indicators of responsiveness. Assessments of ceiling effects include the proportion of patients who score full health within a group of patients with known health detriments. A ceiling or floor effect can affect the sensitivity of the measure in detecting changes over time in patients at the extremes of the measure (for example those with severe disease activity and those with just minor symptoms of the condition).

### 3.2 **Alternative measures** (WP1.2)

No alternative measure searches were performed.

### 3.3 **Evidence required for economic evaluations** (WP1.3)

The existing HTAs were reviewed alongside the variables currently collected in the NCA to determine if clinical or PROM data routinely collected in the NCAs would suffice to address questions of cost-effectiveness, and to identify any gaps in the evidence that would be required to compare providers, or the cost-effectiveness of interventions or policies.

## 4. RESULTS FOR BOWEL CANCER

### 4.1 Evidence of appropriateness of EQ-5D in bowel cancer (WP1.1)

#### *4.1.1 Selection of systematic review*

Two systematic reviews were identified through expert sources.(1;2) The process of selection of the most appropriate review is documented in Table 1.

**Table 1: Selection of most appropriate review for bowel cancer**

| Review | Search date | Relevance of review | Quality of search | Quality of review | Selection |
|---|---|---|---|---|---|
| Oxford (2010)(2) | February 2010 | Question relevant, data detail poor | Reliance on pre-existing database, ovid strategy not provided. However, probably adequate. | QA performed; search numbers provided; unclear if single reviewer DE and SS; synthesis methods unclear | |
| NICEQoL 2014(1) | August 2010 | Question relevant, some data detail available | Searched 7 databases. Supplementary searches in Euroqol database for EQ-5D | QA performed; details of search numbers provided; unclear if single reviewer SS, one reviewer DE; narrative synthesis | Include – more recent than Oxford 2010,(2)more DE detail, more transparent search methods |

QA, quality assessment; SS, study selection; DE, data extraction

#### *4.1.2 Structured abstract for Longworth et al 2014(1)*

Purpose of review

The review aimed to investigate the appropriateness of three generic preference-based health related quality of life (HRQOL) measures (EuroQoL 5 dimensions (EQ-5D), health utility index (HUI-3), and short-form 6 dimension (SF-6D)) for a range of health conditions: vision loss, hearing loss, skin disorder and cancer (bowel, head and neck). This review is only concerned with the results relating to bowel cancer.

Methods of review

Search and study selection: Seven electronic databases were searched from inception: BIOSIS, CINAHL, Cochrane Library, EMBASE, MEDLINE, PsychInfo and Web of Science. Electronic searches were conducted in August 2010. Four sets of search strategies were developed, one for each of the

four health conditions. Search strategies combined terms relating to the health condition with terms for the three HRQOL measures. The full search strategies were listed.

*Inclusion criteria:* Studies were included in the review if they satisfied the following criteria: they contained HRQoL data as measured by one of the three generic HRQoL instruments, namely SF-6D, HU13 or EQ-5D; the study reported results for another HRQoL measure or a disease-specific measure to enable assessment of validity, responsiveness or reliability of the EQ-5D; individuals with one of the four conditions of interest, namely vision disorders, hearing disorders, skin disorders, or cancer. All study types were included.

*Exclusion criteria:* Studies not in the English language.

*Data extraction and synthesis*: Data were extracted by one reviewer using a newly developed standardised form, designed for specific use in the review. Data were tabulated by condition and presented as a narrative synthesis. In addition, results by condition were reported separately according to the construct measured, namely construct validity, reliability and responsiveness. *Construct validity* was measured either using the known-groups methods, or by evaluating convergent validity. The review used the following categories for evidence of correlation: >0.5 was strong; <0.5 to ≥0.3 moderate; <0.3 low. Secondly, responsiveness was assessed where data allowed. Responsiveness was defined as the ability to which an instrument can detect a clinically significant or practically important change in health utility over time.

**Authors' conclusions:** The authors of Longworth et al. 2014a concluded that the evidence for the appropriateness of the EQ-5D was generally satisfactory, with positive evidence of the construct validity of the EQ-5D but insufficient evidence to support the responsiveness of the EQ-5D for this health condition.

#### *4.1.3 Assessment of the review in relation to objectives of work package 1.1*

Relevance of review question: The aim of Longworth et al. 2014 was to assess whether the three HRQoL measures were appropriate for use for four specific conditions. Whilst only one subgroup of one of these conditions is relevant to the aims of WP 1.1 (specifically bowel cancer), the inclusion criteria and methods used for assessment of these relevant conditions are concordant with WP 1.1. Only data relating to bowel cancer will be reported in the following section.

Assessment of review quality: The assessment was conducted using a modified version of the 'assessing the methodological quality of systematic reviews' (AMSTAR) tool (3) and also by considering the strength and quantity of the evidence. The adequacy of the reported data in the context of work package 1.1 was also assessed. A summary of the quality assessment is shown in Table A7.

Longworth et al. (1) scored well against most of the relevant AMSTAR criteria. A full study protocol is published with the review to evidence an a priori design, therefore reducing the possibility of the existence of reporting bias. Quality assessment of the included studies was conducted however no formal tool was used for this purpose. The authors state that a 'judgment regarding the risk of bias of each study was determined by reviewing methods of patient recruitment, and noting any missing data reported (either study drop-outs or incomplete questionnaires)' (1). Study selection was carried out by only one reviewer, and double data extraction or data-checking was not conducted, leaving the study at higher risk of errors. Inclusion criteria are clearly defined, however the only exclusion criterion provided is that of papers in non-English languages.

Acceptability of the search: Unlike the other reviews, an iterative search approach was applied in the review, although this was not fully described. A comprehensive Medline strategy comprising generic instrument search terms and condition terms were reported in the review. The search is considered comprehensive for the purpose of the review.

Acceptability of study selection: Study selection criteria were clearly defined and concordant with the inclusion criteria for WP1.1. Whilst the review covered four health conditions, separate analyses were conducted for each and therefore study selection was acceptable.

Adequacy of available data and synthesis: The review provided sufficient data on each of the properties of interest and was therefore adequate for the requirements of WP1.1.

In conclusion, the methods employed in the review were generally of an acceptable quality and designed to meet the requirements of WP1.1. However, the study search was conducted in 2010, and as such an update search was conducted, and any new studies integrated with the findings of Longworth et al.(1)

A total of 98 studies were included in the Longworth review, reporting results for 20 different types of cancer. The authors report 11 of these studies evaluated the construct validity or responsiveness of the EQ-5D, SF-6D or HUI-3 for colon cancer.(4-14) One study reports data for HUI3 only (i.e. no EQ-5D data),(12) and one study is reported in study characteristics only, with no associated outcome data so is not included here.(13)

From a total of 76 titles identified in the update search, no additional studies met the inclusion criteria for WP1.1. Therefore a total of nine studies with data relating to EQ-5D are included in this review. Of the nine included studies, one study is explicitly reported to have used the UK EQ-5D tariff (10). No further details of EQ-5D versions are provided for the remaining studies. Three studies were conducted in the UK (4;9;10). The majority of the remaining studies were conducted in other European countries, one in Sweden;(8) and three in the Netherlands.(5;6;11) One study was conducted in Japan (7), and one study was multinational.(14)

Details of patient characteristics were not reported in great detail in the review, for example no details of sex or age were given. All studies included patients with colorectal cancer. In seven studies patients were undergoing or had recently undergone surgery (4-9;11), whilst in two studies patients were assigned to a pharmaceutical intervention group. (10;14)

Three studies were RCTs (8;10;14), two studies were cross-sectional,(6;11) and four were before/after studies.(4;7;9;11)

A range of measures were used to assess the construct validity and/or responsiveness of the EQ-5D. Some measures used for comparison were designed for generic cancer use: European Organisation for Reseach and Treatment of cancer (EORTC)(6;8;14), or EORTC-QLQ.(5) Two studies used measures designed for use in generic cancer therapy functional assessment of cancer therapy – emotional well-being, and colorectal cancer (FACT-EW),(4) FACT-C(9). Doornebosch et al.(11) compared the EQ-5D with a condition specific measure, the Faecal Incontinence Quality of Life (FIQL), and Siena et al.(14) used the FACT Colorectal Symptom Index (FCSI). Wilson et al. compared the EQ-5D with other generic HRQoL measures (SF-12 and EORTC QLQ),(9) whilst Sharma et al. assessed the EQ-5D in relation to tools designed to assess mood, Positive and Negative Syndrome Scale (PANSS), and Hospital Anxiety and Depression Scale HADS.(4) The evidence is presented in accordance with the definitions of psychometric properties given in Section 3.1.2.

**Construct validity (known group):** Six studies were identified in the review that evaluated the construct validity of the EQ-5D in bowel cancer patients using the known-group method.(5;6;8;9;11) Positive evidence for the EQ-5D was reported in two studies.(9;11) The study by Wilson (2006) showed that the EQ-5D was consistent with preoperative EQ-VAS, SF-12 general health, SF-12 PCS FACT-C total score and Quality of Life Questionaire (QLQ) general health, with all measures showing a significant difference in scores between Easten Cooperative Oncology Group (ECOG) performance status groups. Doornebosch (2008) showed that both EQ-5D and FIQL scores were not affected by age and gender of the patients, nor by surgical aspects or tumour characteristics. Two further studies presented negative evidence for the EQ-5D.(5;6) Doornebosch et al. 2007 reported that the EORTC QLQ-CR38 (the colorectal module of the EORTC) showed a significant difference in scores by surgical intervention groups for defecation problems, but neither the EQ-5D nor the EORTC QLQ-C30 detected a statistically significant difference in the overall scores (EQ-5D) or any subscale (EORTC QLQ-C30).(5) Gosselink et al. studied the difference in EQ-5D and EORTC scores for three surgical intervention groups: Colo-anal J-pouch anastomosis (CPA), abdominoperineal resection (APR), and transanally double-stapled low colo-rectal anastomosis (LRA). EQ-5D scores did not differ between the three surgical intervention groups, whilst scores on four subscales of the EORTC (global health status in the QLQ-C30, and body image, micturition and defacation in the QLQ-CR38) did differ significantly in some between-group comparisons.(6) Gosselink et al. also presented some mixed or equivocal evidence for the construct validity of the EQ-5D, along with two other studies.(6-8) Gosselink et al. found that mean scores on the EQ-5D differed for one of the three surgical interventions (CPA), when compared to a sex-age matched general population. Hamashima et al. failed to show significant differences on EQ-5D scores between patient groups (with or without stoma), though it was unclear if a difference should be expected.(7) Janson et al. reported that the EQ-5D did not detect a difference between study groups at baseline, in agreement with findings as measured by the EORTC-C30.(8)

**Construct validity (convergent):** One study reported results for convergent validity.(4) The EQ-5D index score was not significantly correlation with TNM stages (r=0.06, p=0.5), where the HADS anxiety score, positive and negative affect schedule (PANAS) negative affect score and FACTC-emotional wellbeing score showed statistically significant low to moderate correlations (r= 0.345, r=0.294, r=-0.354 respectively, p=0.01 for all). However, the HADS depression, PANAS positive affect, MRS, FACTC-physical wellbeing, FACTC social and family wellbeing, FACTC functional wellbeing and

FACTC- colorectal cancer-specific scores were all non-significantly correlated to TNM (Tumour, Node, Metastases) stage as well.

**Responsiveness:** Five studies were identified in the review that assessed the responsiveness of the EQ-5D.(4;8;10;11;14) One study presented positive evidence for the EQ-5D: in an RCT of pharmaceutical interventions, Anderson and Palmer (1998) showed statistically significant changes in EQ-5D scores 2 weeks post baseline in favour of patients in the raltitrexed group, showing improvements for mobility and usual activities, and for general health. Statistically significantly different changes were reported at 2 weeks in the raltitrexed and the 5-FU+LV (fluorouracil + leucovorin) groups using the Rotterdam Symptom Checklist (RSCL). These results were no longer significant by week 10, although the EQ-5D did show non-significant trends in favour of the raltitrexed group. Two studies reported a lack of responsiveness in the EQ-5D.(4;11) Doornebosch et al. reported that significant changes were not detected by the EQ-5D where they were detected by the overall Faecal Incontinence Severity Index (FISI), and two of the four dimensions of the Faecal Incontinence Quality of Life (FIQL) measure (lifestyle and embarrassment, but not coping and depression) 6 months after surgery.(11) Sharma 2007 explored the validity of the EQ-5D in patients undergoing elective open resection. EQ-5D scores indicated no significant change over time (in keeping with HADS anxiety, both PANAS positive and negative affect scores, mood rating scale, FACTC total and trial outcome idex and EQ-VAS) where the HADS depression score did.(4) Two studies reported data that was either mixed or equivocal about the responsiveness of the EQ-5D.(8;14) Janson et al. (2007) found that in general, the differences in mean scores over time in the EQ-5D data follow the same trend as observed in the majority of EORTC symptom and function scales with no significant differences reported for either. However, the EQ-5D did not accurately detect the very small statistically significant changes in the EORTC scales observed by the emotional and social scales, between baseline and week 4.(8) In Siena et al. the EQ-5D was only able to detect a statistically significant difference in the treated group whereas improvements in symptoms measured by the FACT colorectal symptom index were observed in both the treated and best supportive care groups.(14)

### *4.1.4 Conclusion of appropriateness of EQ-5D in bowel cancer*

The evidence base assessing the performance of the EQ-5D in bowel cancer is of a moderate size (N=9), but not all studies assessed all the psychometric properties of the EQ-5D. With the exception of one study which used UK preference-based weights,(10) it is unclear which tariff was used. No details were provided in Longworth et al. on characteristics such as age.(1)

No evidence was reported in the review for either acceptability or reliability of the EQ-5D. Construct validity was explored using known group methods in six studies, and convergent methods in one study. The evidence relating to construct validity by known group methods was mixed, with two studies reporting that the EQ-5D was able to detect differences between groups where other measures did,(9;11) two studies reporting that the EQ-5D failed to detect a difference in groups where other measures did,(5;6) one study reporting no differences between groups by the EQ-5D and EORTC QLQ-C30(8) and one study reporting no difference in EQ-5D scores between two groups where it was not clear if a difference should be expected.(7) The evidence-base relating to convergent validity was small with only one study. This study reported the EQ-5D was not correlated with TNM stage where scores from three other measures or subscales were moderately correlated, and six other measures or subscales weren't statistically significantly correlated.(4) However, the sample size (n=104) may have resulted in under-powering when split across several TNM stages. Results relating to responsiveness were also mixed, with one study showing the EQ-5D was able to detect a change where another measure did,(10) two studies showing the EQ-5D was not able to detect a change where other measures did,(4;11) and two studies showing mixed results where the EQ-5D was able to detect some changes but not others.(8;14)

In summary, whilst there is some strong positive evidence to support the use of the EQ-5D in this patient group, the negative and mixed evidence suggest that additional validation is required before the EQ-5D can be recommended (Table 2).

**Table 2: Summary of evidence on EQ-5D for bowel cancer**

| Measure (N) | Acceptability | Reliability | Construct (KGV; Convergent) | Responsiveness (Change over time; Ceiling effects) |
|---|---|---|---|---|
| Adults | | | | |
| EQ-5D (9) | Not reported | Not reported | Mixed; Poor (n=1) | Mixed; Not reported |
| EQ-5D requires additional validation. | | | | |

### 4.2 Alternative measures in bowel cancer (WP1.2)

Whilst the EQ-5D is considered to require further validation in patients with bowel cancer, alternative measures were not reviewed. It is recommended that a cancer specific PROM, the European Organization Quality of Life Questionnaire (EORTC QLQ-C30) and it's relevant bowel-specific module EORTC CR29 (which supersedes the CR38) is collected alongside the EQ-5D as

condition specific measures may be more sensitive to the effects of interventions on the condition specific symptoms and the side effects of treatments.

The EORTC QLQ-C30 consists of 30 questions covering function (e.g. cognitive, emotional, physical, role, social) and the common cancer symptoms (e.g. fatigue, nausea and vomiting, pain).(15) Responses to these are summarised using 14 sub-scales plus a global quality of life scale.(15) A recently developed UK based preference-based utility tariff can be used to generate utility values in economic evaluations.(16) However, it should be noted that the the utility values obtained using this tariff are not directly comparable to those generated using the EQ-5D.

The EORTC QLQ-CR38(17)was developed in the Netherlands, and comprises 38 items. Nineteen questions are completed by all patients and the other 19 by subgroups of patients such as those with or without a stoma. The measure comprises two functional scales (body image and sexuality) and seven symptom scales (micturition problems, symptoms in the area of the gastrointestinal tract, chemotherapy side-effects, problems with defaecation, stoma-related problems, male and female sexual problems). Whilst there is some psychometric data relating to it, it was never fully validated.(18) It has been superseded by the CR29, which was developed with patients from the UK, France and Germany (19) and has been validated in at least one study,(18) and adapted for use in at least two other countries (Spain and Iran).(20;21) It comprises 29 items relating to colorectal cancer and its treatment, including micturition, pain, defecation problems, faecal incontinence, anxiety, body image, abdominal bloating, dry mouth, hair loss, taste problems, skin problems, stoma embarrassment and stoma problems. Gender-specific sexual items were also included.

### 4.3 **Evidence for economic evaluations in bowel cancer** (WP1.3)

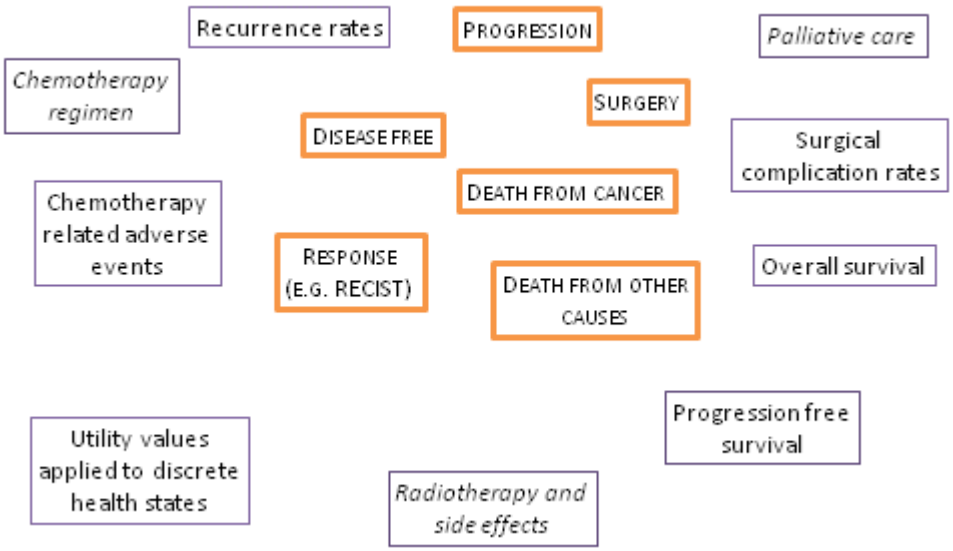
#### *4.3.1 Cost-effectiveness modelling approach used in recent HTAs in bowel cancer*

Ten TAs relating to bowel cancer were identified from the searches. Four of the TAs were superseded by more recent publications,(22-25) leaving three MTAs,(26-28) and three STAs.(29-31) A clinical guideline (CG) was subsequently identified from the references lists of the included studies.(32) Six of the TAs examined the clinical and cost-effectiveness of pre-specified pharmaceutical interventions and one examined the clinical and cost-effectiveness of laparoscopic surgery compared to open-resection (Table 3).(27) While the CG incorporated a broader decision area covering both the diagnosis and management of colorectal cancer, their economic model

focussed on the costs and effects associated with ten sequences of chemotherapy limited to two lines of treatment.(32)

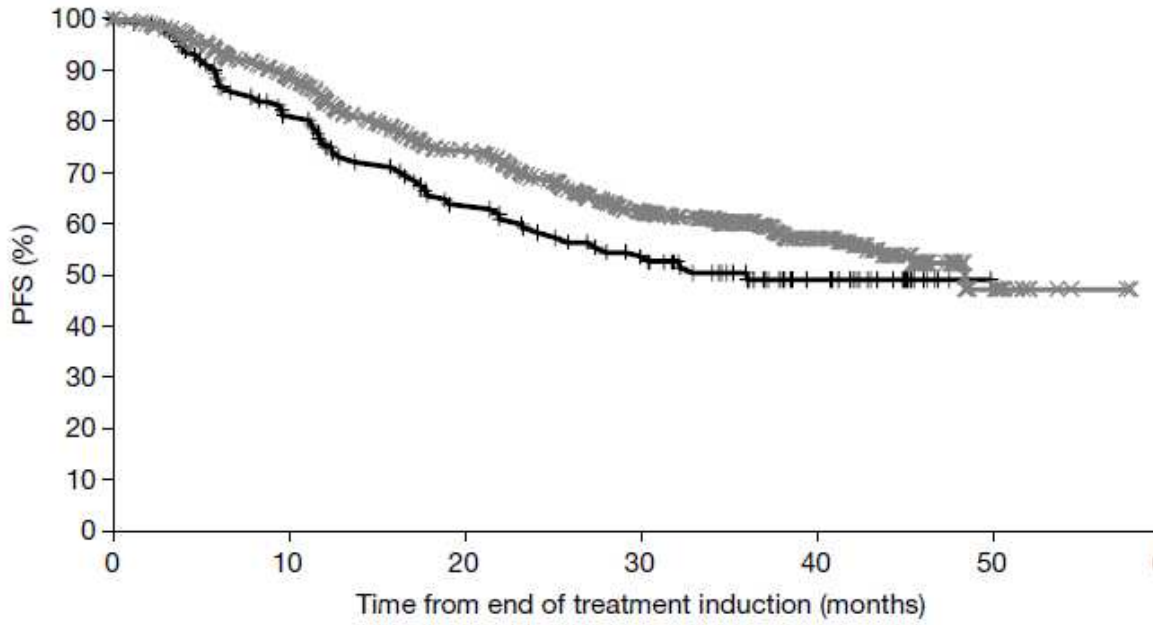
With the exception of the CG, which used a decision tree,(32) state transition models were used to examine the cost-effectiveness of the interventions under appraisal. All seven models comprised of discrete health states which represented the clinical pathway for people with bowel cancer at the point of the intervention. The number of health states was typically less than five and the health states were predominantly defined in terms of progression or relapse (worsening of cancer symptoms) and survival (Figure 1). Clinical trial data (survival curves and hazard ratios (HR), see exemplar provided in Figure 2) were used to inform overall survival (OS), tumour response rates (e.g. defined using the RECIST instrument) and progression free survival (PFS). Treatment specific rates for graded adverse events were also sourced from the clinical trials used to inform the effectiveness of the interventions. The TA for laparoscopic surgery used survival curves from RCTs to model OS, surgical mortality rates, and recurrence. Intervention specific rates for non-elective surgery and hernias were sourced from the literature.

**Figure 1: Modelling approach used in bowel cancer HTAs**




Legend: Orange framed boxes with uppercase text describe the health states used in the diabetes TA models while the purple framed boxes with lower case text describe the evidence used. Italised text indicative of additional variables which would be informative for future economic evaluations.

**Figure 2: Exemplar survival curves used to model interventions in cancer**




PFS=Progression free survival

All studies quality adjusted survival by assigning mean utility values to the discrete health states. With the exception of one appraisal which utilised data from an observational study conducted to inform the TA,(31) the utilities were either drawn from the study used to provide the primary clinical evidence,(26;30) or were sourced from the literature.(27-29;32) Rather than modelling the effects on HRQoL explicitly in separate health states, adjustments to utilities were applied in some models to reflect the prevalent side-effects of treatments for cancer (e.g. skin-related toxicities, hypomagnesaemia, paronychia, abdominal pain, diarrhoea and constipation, fatigue) or adverse events (e.g. surgical complications).(26;29-32) None of the studies used EQ-5D data for all the utilities within the models and many used utility values which were not weighted by general population preferences. The results of the searches conducted to inform the model parameters suggest the volume of EQ-5D data in patients with colorectal cancer is very limited with many of the authors recommending this as a future research priority. It has also been suggested that evidence categorised by Duke's stage would be useful. Staging is currently used for chemotherapy licensing indications and would be useful when modelling the cost-effectiveness of screening interventions; it could be linked to audit data to estimate resource usage by stage; and if stage was linked to EQ-5D utility values this would also be useful for the future economic models in bowel cancer as this evidence is currently not available.[personal communication Dr P Tappenden, ScHARR 3rd June 2014]

**Table 3: Summary of existing models used in bowel cancer TAs**

| Model method, clinical effect | Method used to model utilities |
|---|---|
| MTA (TA242): Colorectal cancer (metastatic): 3rd line, cetuximab, bevacizumab and panitumumab; 2012(26) | |
| TAG State tranisition model<br>3 discrete health states: PFS, progressive disease, death (progression defined as worsening symptoms of cancer)<br>Effectiveness: survival curves for OS, response (using Dukes stage) and PFS<br>Source: RCTs used for clinical effect | Utility: HUI3; mean values assigned to discrete HS<br>Source: RCT used for clinical effect<br>AEs: intervention specific utilities used to account for differences in treatment toxicity |
| CG (CG131): Colorectal cancer: the diagnosis and management of colorectal cancer; 2011(32) | |
| Decision tree<br>3 discrete health states: stable metastatic disease, progressive metastatic disease, death<br>Effectiveness: survival curves for OS, response and PFS<br>Source: RCTs used for clinical effect | Utility: elicited (TTO) from patients and community members; mean values assigned to discrete HS<br>Source: published literature<br>AEs: data from patients with metastatic breast cancer used to account for treatment toxicities such as diarrhoea/vomiting, hand foot syndrome (no details on measure/elicitation method) |
| STA (TA212): Colorectal cancer (metastatic): bevacizumab; 2010(29) | |
| State transition model<br>4 discrete health states: 1st line treatment, PFS post treatment, progression, death<br>Effectiveness: survival curves for OS, response and PFS<br>Source: RCTs used for clinical effect | Utility: EQ-5D, HUI, supplemented with expert opinion; mean values assigned to discrete HS<br>Source: published literature<br>AEs: assumed higher utility when not on treatment to account for intervention toxicity |
| STA(TA176): Colorectal cancer (first line) – cetuximab; 2009(30) | |
| State transiition model<br>Numerous discrete health states (for each of 1st to 3rd line therapy): successful resection, unsuccessful curative resection (i.e. progression), death<br>Effectiveness: survival curves for OS, response and PFS<br>Source: RCTs used for clinical effect | Utility: EQ-5D, HUI, assumption; mean values assigned to discrete HS<br>Source: studies used for clinical effect, supplemented by published literature<br>AEs: disutilities assumed to be captured in main data as from patients receiving first line treatment |
| STA (ID514): Colorectal cancer (metastatic) - aflibercept (31) | |
| State transition model<br>4 discrete health states: no progression on treatment, no progression post treatment, progressive disease, death<br>Effectiveness: survival curves for OS, response and PFS<br>Source: RCTs used for clinical effect | Utility: EQ-5D; mean values assigned to discrete HS<br>Source: observational utility study, supplemented by published literature<br>AEs: disutilities assigned using published data for treatment toxicity |
| MTA(TA105): Colorectal cancer - laparoscopic surgery; 2006(27) | |
| TAG semi-Markov model<br>5 discrete health states: disease free, treatable (surgical or other) recurrence of disease, disease free after recurrence, non-operable recurrence, death<br>Effectiveness: survival curves for OS, surgical mortality rates, recurrence; rates for non-elective surgery and hernias<br>Source: survival curves from RCTs used for clinical effect; rates sourced from literature | Utility: EQ-5D, elicited from oncology nurses (method unclear); mean values assigned to discrete HS<br>Source: published literature<br>AEs: no mention of disutility due to adverse events |
| MTA(TA100): Colon cancer (adjuvant) - capecitabine and oxaliplatin; 2006(33) | |

| | |
|---|---|
| TAG State transition model<br>3 discrete health states: alive without relapse, alive post-relapse, death<br>Effectiveness: survival curves for OS, response and PFS<br>Source: RCTs used for clinical effect | Utility: elicited from patients with colorectal cancer using SG; mean values assigned to discrete HS<br>Source: published literature<br>AEs: no mention of disutility due to treatment toxicity |

HS: health states; AE: Adverse Events; MTA: Multiple Technology Appraisal; STA: Single Technology Appraisal; CG: Clinical Guideline; TAG: Technology Appraisal Group; TA: Technology Appraisal; TTO: Time trade-off; SG: Standard Gamble; RCT: randomised controlled trial; OS: overall survival; PFS: progression free survival

In summary, the following evidence would be required to compare providers or the cost-effectiveness of interventions for bowel cancer:

- Condition severity (e.g. Dukes stage)
- Tumour location
- Surgical rates (type of intervention, success rate, post-surgical complication, length of stay)
- Chemotherapy regimens (medications, adverse events)
- Radiotherapy (type of intervention, success rates, side-effects)
- Recurrence/relapse rates (with dates)
- Utility values
- Death with cause to model cancer related deaths

The majority of this evidence would need to be dated and linked through timings of collection.

### *4.3.2 Fields collected in the bowel cancer NCA*

All patients[1] with a diagnosis of bowel cancer admitted for the first time to a NHS Trust in England or a Health Board in Wales are eligible for inclusion in the bowel cancer audit. The fields in the bowel cancer NCA are collected via an excel spreadsheet (Bowel_Dataset_v1.3) completed by NHS staff (see Appendix). The data provide information on patient characteristics (age, gender, height, weight); tumour and classification (primary site, Dukes' staging, metastases); dates, outcomes and complications associated with procedures (barium enema, colonoscopy, CT and MRI scans, ultrasounds) and interventions (elective and scheduled surgery, radiotherapy, chemotherapy, palliative care); complications of cancer, follow-up care (clinical status, recurrence of primary

[1] The bowel cancer NCA is likely to include a very small proportion of paediatrics, but for the purpose of the current study, this audit is considered to be in adults only.

tumour and metastasis spread, treatment morbidity); and survival status (date and morbidity code). There does not appear to be a patient questionnaire in the bowel cancer NCA.

### *4.3.3 Comparing fields in bowel cancer NCA with variables used in existing HTAs*

The existing models in bowel cancer use survival curves for recurrence, progression and OS. The NCA information on clinical interventions (tumour, treatment, follow-up) would provide some of the information required to compare alternative treatments. The mortality date could be used to model overall mortality, and there may be sufficient detail to extract survival curves for progression and recurrence from the mandatory fields. Side-effects and adverse events due to chemotherapy, radiotherapy and surgery are prevalent. While there is some information on toxicity (treatment related morbidity: mild toxicity, moderate toxicity, severe toxicity, death due to toxicity) this field is non-mandatory (Table A7) and it is not clear if the level of detail collected would suffice to populate a model.

Clinical variables such as American Society of Anesthesiologists classification (ASA) grade (a normal healthy patient, a patient with mild systemic disease, a patient with severe systemic disease, a patient with severe systemic disease that is a constant threat to life, a moribund patient who is not expected to survive without the operation, a declared brain-dead patient whose organs are being removed for donor purposes, not known) and Dukes' stage (A, B, C1, C2, D, not known) might be used to case-mix patients when comparing performance or cost-effectiveness of interventions. However, Duke's staging requires pathologic confirmation and it is not clear if this would be available for all patients in the audit.

Patient related outcome measures are not currently collected in the bowel cancer NCA. The inclusion of a preference-based HRQoL questionnaire (e.g. the EQ-5D), would be extremely informative, particularly as the existing HTAs do not in generally use preference-based data to weight survival due to a dearth of appropriate evidence in patients with bowel cancer. As stated earlier, the use of Duke's staging alongside the preference-based measure would be extremely useful evidence for use in economic models. This approach would likely require an analysis to link the two variables but could potentially reduce the uncertainty in economic results. There does not appear to be any suitable variable collected in the bowel cancer audit which might be used to generate proxy preference-based utility values using an existing function.

Assuming the mandatory fields have relatively high completion rates, with the exception of HRQoL, and toxicity, the information currently collected in the existing NCA would provide the majority of information required to model the cost-effectiveness of interventions and policies in bowel cancer. As previously noted, there is a dearth of preference-based HRQoL evidence in bowel cancer, and the collection of utilities within the NCA would be recommended as an important and valuable consideration. It is not known if there are any ongoing studies in this area, but the bowel cancer NCA will be undergoing a retendering process later this year (2014) and it is possible that the new contract will include the collection of PROMS.[personal communication, Eleanor Bunn, Audit Coordinator, 13th May 2014]

### 4.4 Recommendations for bowel cancer

The evidence base relating to the appropriateness of the EQ-5D in patients with bowel cancer is mixed and further validation is required. All the evidence is in adults and while the NCA inclusion definition includes all patients with bowel cancer, the proportion of paediatrics is likely to be extremely small and is not considered in this section. The evidence in the literature which could be used to populate the utility values in economic models is extremely poor (see Section 4.3). With the exception of information on treatment related adverse events/complications, it is thought that the current NCA collects much of the information required to conduct economic evaluations. However, as far as we are aware, there does not appear to be a patient questionnaire. Potential recommendations (PR) and areas for future research (FR) are discussed below. All suggested future research areas are indicative and would require a discussion and detailed proposal if required.

Due to the extremely limited evidence base providing information of HRQoL in this patient group, it is recommended that consideration is given to including a preference-based measure in the NCA for bowel cancer. This would involve developing a patient questionnaire suitable for patients with bowel cancer. While it is recommended that the EQ-5D is considered in the first instance (PR.1), due to the equivocal evidence on the psychometric properties of this instrument in this patient population, it is also recommended that a psychometric assessment of the EQ-5D is conducted on the initial round of data collected (FR.1). The inclusion of a cancer specific measure such as the EORTC QLQ-C30 and the associated colorectal module (QLQ-CR29) is also recommended (PR.2), to identify cancer specific issues which will be useful when comparing aross providers, and to inform future economic evaluations.

To assist in future economic evaluations, it is also recommended that consideration is given to ensuring the mandatory fields in the audit include some form of measure to assess the stage of the cancer (e.g. Duke's staging) (PR.2). This should be synchronised with the timing of collection of the HRQoL data. In addition, the majority of patients will receive chemotherapy and the mandatory collection of treatment related adverse events is recommended (PR.3). A thorough detailed inventory of the exact information collected in the NCA is required before recommendations relating to additional mandatory fields are suggested (PR4, FR.3).

**Table 4: Recommendations and associated future research for bowel cancer**

| | |
|---|---|
| ***PR.1*** | *Include the EQ-5D in future patient questionnaires alongside a condition specific measure such as the EORTC QLQ-C30 and the colorectal module (QLQ-CR29)* |
| ***FR.1*** | *Assess the psychometric properties of the EQ-5D using the data collected in the bowel cancer audit* |
| ***PR.2*** | *Include a severity measure such as Dukes' staging (to be collected alongside the HRQoL data)* |
| ***PR.3*** | *Collect mandatory information on adverse events associated with chemotherapy regimens (plus radiotherapy adverse events, and surgical complications)* |
| ***PR.4*** | *Include additional mandatory fields in the bowel cancer audit* |
| ***FR.3*** | *Detailed analyses of fields currently collected in the bowel cancer audit to identify recommendations for future mandatory fields* |

## 5. SUMMARY

The following section provides an overview of the results presented within the individual sections of the report. A summary of the evidence used to inform the conclusions for WP1.1 and WP1.2 is provided in Table 5. This section provides an overview only and it is recommended that the preceding sections are used for details on particular conditions.

### 5.1 Summary of evidence used to inform the conclusions for WP1.1 and WP1.2

A reanalysis and update of an existing review (n=9 primary studies) provided mixed evidence regarding the construct validity and responsiveness of the EQ-5D. Acceptability, reliability and ceiling effects were not reported by the original review authors. In some cases, known group validity appeared good against both cancer specific and generic measures, whilst in others it was unable to detect differences between groups where condition specific measures (EORTC QLQ-C30 or QLQ-CR38) could. One study reported convergent validity of the EQ-5D against TNM stages was low, as was the case for several other cancer specific and psychological symptom specific measures or subscales. However, some measures and subscales had moderate correlations. Responsiveness was also mixed, with the EQ-5D able to detect a change in health status over time in some cases, but not in others, where condition specific or symptom specific measures and subscales such as the HADS depression score, FACTC and EORTC measures did. It was concluded that the EQ-5D could not be recommended without further validation. A review of all alternative measures that could be used was not conducted, but it is recommended that the EQ-5D be used in conjunction with the EORTC-C30, which has a UK utility tariff, and the colorectal module EORTC QLQ-CR29.

**Table 5: Summary of evidence supporting the psychometric properties of EQ-5D in all conditions**

| Measure | N | Acceptability | Reliability | Construct | | Responsiveness | | Overall |
|---|---|---|---|---|---|---|---|---|
| | | | | KGV | Convergent | Change over time | Ceiling Effect | |
| EQ-5D | 9 | NR | NR | Mixed | Poor | Mixed | NR | Acceptable |
| EORTC QLQ-C30 EORTC QLQ-CR38/29 | | The psychometric properties of these measures have not been reviewed in the current report | | | | | | |

N= number of studies used to inform conclusions, KGV: known group validity; NR, the existing review did not review this psychometric property;

### 5.2 **Summary of evidence required for use in economic evaluations** (WP1.3)

The bowel cancer audit does not include a patient questionnaire thus PROMs are not currently collected and existing literature on preference-based data which could be used to inform formal economic evaluations is sparse. The audit collects some of the information required to compare providers and economic evaluations such as survival, progression and staging of disease (Dukes' stage). However, key variables such as toxicity due to chemotherapy and adverse effects of surgery and radiotherapy are not currently mandatory fields.

## Appendix

The tables in this Appendix provide additional information for the reviews (WP1.1, 1.2 and 1.3) conducted for bowel cancer.

**Table A1: Quality assessment of Longworth et al 2014 for bowel cancer(34)**

| Quality assessment criteria | Compliance with criteria |
|---|---|
| **AMSTAR** | |
| Was an a priori design provided? | Yes |
| Was there duplicate study selection and data extraction? | No |
| Were the methods used to combine the findings of the studies appropriate? | Unclear - narrative synthesis conducted but no justification for lack of meta-analysis provided. |
| Was the scientific quality of the included studies assessed and documented? | Yes, but not by validated method. |
| Was the scientific quality of the included studies used appropriately in formulating conclusions? | Yes – studies not excluded for quality reasons but quality assessment stated to be used for informing strength of evidence. |
| **Overall judgement of quality of review** | Good but only 1 reviewer. |
| **Quality of the searches** | Acceptable |
| **Strength of the evidence** | |
| Were the conclusions robust and conclusive? | No, conclusions were robust and conclusive for construct validity, but only one study was available with responsiveness data. |
| **Quantity of the evidence** | |
| Was there enough data to be confident that any additional data published subsequently would be very unlikely to change the conclusions drawn? | No – only one study for responsiveness data. |
| **Adequacy of data reported** | |
| Did the review provide sufficient data to allow integration of an update/assessment of the methods used? | Yes |
| Did the review assess EQ-5D in a way compatible with the aims of work package 1.1? | Yes, construct validity (known groups or convergent) or responsiveness (effect sizes, standardised response means, or correlation with change scores on symptom measures). |

**Table A2: Characteristics of primary studies included in Longworth review (bowel cancer). Adapted from Longworth et al.(34)**

| Author, year | Sample size | Condition | Study design | Study information | Male/female (%) | Mean age at baseline |
|---|---|---|---|---|---|---|
| Anderson and Palmer, 1998(10) | 545 | Advanced colorectal cancer | RCT | Raltitrexed vs. Standard 5-fluorouracil | N/R | N/R |
| Doornebosch et al, 2007(5) | 62 | T1 carcinoma after surgery (TEM), T1 to T3 (35%)(TME) | Cross-sectional | Total Mesorectal excision vs. Transanal Endoscopic microsurgery | N/R | N/R |
| Doornebosch et al., 2008(11) | 47 | People with rectal cancer eligible for TEM | Before-after | Transanal Endoscopic microsurgery | N/R | N/R |
| Gosselink et al., 2006(6) | 204 | People with rectal cancer in the middle or low third of the rectum after total mesorectal excision | Cross-sectional | Abdominoperineal resection; Transanally double stapled low colorectal anastomosis; colo-anal J-pouch anastomosis | N/R | N/R |
| Hamashima, 2002(7) | 110 | Rectal cancer patients who had received surgery as their initial treatment | Before-after | Surgery | N/R | N/R |
| Janson et al., 2007(8) | 285 | Elective colon cancer patients with potentially curable cancer best treated by right or left hemicolectomy or sigmoid resection | RCT | Laparoscopic colon resection vs. open resection | N/R | N/R |
| Sharma et al., 2007 (4) | 104 | Newly diagnosed colorectal cancer scheduled for elective open resection | Before-after | Elective open resection | N/R | N/R |
| Siena et al., 2007(14) | 463 | Metastatic colorectal cancer patients who had progressed on prior fluoropyrimidine, irinotecan and oxaliplatin | RCT | Panitumumab plus best supportive care vs. best supportive care alone | N/R | N/R |
| Wilson et al., 2006(9) | 210 | Patients undergoing potentially curable open | Before-after | Surgery | N/R | N/R |

| Author, year | Sample size | Condition | Study design | Study information | Male/female (%) | Mean age at baseline |
|---|---|---|---|---|---|---|
| | | surgery for colorectal cancer | | | | |

RCT: randomised controlled trial; TEM: transanal endoscopic microsurgery; TME: total mesorectal excision;

**Table A3: Characteristics of primary studies included in Longworth review (bowel cancer). Adapted from Longworth et al.(34)**

| Author, Year, Location | EQ-5D | Comparison measure | Psychometric properties assessed | Assessment of psychometric properties |
|---|---|---|---|---|
| Anderson and Palmer, 1998, UK(10) | EQ-5D UK | RSCL | Responsiveness | Odds ratio for responses of EQ-5D dimensions between baseline and week 5 and 15 over the two groups |
| Doornebosch et al, 2007, Netherlands(5) | EQ-5D, EQ-VAS | EORTC QLQ | Known group validity (severity) | EQ-5D scores according to intervention group |
| Doornebosch et al., 2008, Netherlands(11) | EQ-5D | FIQL | Responsiveness | Wilcoxon's signed rank test and Mann-Whitney U test for change scores within or between groups. Spearman rank correlations between change scores. |

EQ-5D: Euro-QoL 5 dimensions; RSCL: Rotterdam symptom checklist; VAS: visual analogue scale; EORTC QLQ: European organisation for research and treatment of cancer core quality of life questionnaire; FIQL: faecal incontinence quality of life.

**Table A4: Construct validity results for bowel cancer, adapted from Longworth et al 2014.(34)**

| Author, year | Method of measuring known groups validity | Known group validity results |
|---|---|---|
| Construct validity (known group) | | |
| Doornebosh 2007(5) | EQ-5D, EORTC QLQ C30 and EORTC QLQ-CR38 scores according to intervention group | No difference on EQ-5D scores between the three intervention groups.<br>No difference in EORTC QLQ-C30 subscales scores between the three intervention groups.<br>Significant difference in scores for EORTC QLQ-CR38 between TEM and TME groups regarding defecation problems – TEM patients had less defecation problems than TME patients ($p<0.05$). |
| Gosselink 2005(6) | EQ-5D and EORTC scores according to intervention group | Mean EQ-5D score for CPA group was significantly higher than the sex-age matched general population.<br>Mean EQ-5D scores for LRA and APR groups were similar than the sex-age matched general population.<br>EQ-5D scores did not differ between the three intervention groups.<br>5 subscales of EORTC scores were significantly different between the three intervention groups . |
| Hamashima 2002(7) | Significant differences in EQ-5D scores between patients with or without stoma | EQ-5D scores for those with and without stoma groups were not significantly different. |
| Janson 2007(8) | EQ-5D and EORTC scores according to intervention group | EQ-5D scores at baseline between intervention groups were not significantly different.<br>EQ-VAS scores at baseline between intervention groups were not significantly different.<br>EORTC QLQ-30 scores at baseline between intervention groups were not significantly different. |
| Wilson 2006(9) | EQ-5D scores by ECOG performance status groups<br>EQ-VAS, SF-12 GH, SF-12 PCS, QLQ GH, FACT-C scores by ECOG performance status groups | EQ-5D, scores were significantly different between ECOG PS status groups.<br>EQ-VAS, SF-12 GH, SF-12 PCS, QLQ GH were significantly different between ECOG performance status groups.<br>EQ-5D, total scores declined with advancing preoperative ECOG performance status.<br>EQ-VAS, SF-12 GH, SF-12 PCS, QLQ-GH, FACT-C |
| Doornebosch 2008(11) | EQ-5D scores vs FISI scores and FIQL scores pre and 6 months post surgery. | EQ-5D scores were not affected by age and gender of the patients, surgical aspects and tumour characteristics.<br>FIQL scores were not affected by age and gender of the patients, and surgical aspects and tumour characteristics. |
| Construct validity (convergent) | | |
| Sharma 2007(4) | Correlations between measures and TNM stage | EQ-5D correlation with TNM stages is small and not statistically significant.<br>( HADS anxiety, PANSS negative affect and FACT-emotional wellbeing were all significantly moderate correlated with TNM stage) |

EQ-5D: Euro-Qol 5 dimensions; EORTC QLQ: European organisation for research and treatment of cancer core quality of life questionnaire; TEM: transanal endoscopic surgery; TME: total mesorectal excision; CPA: colo-anal J-pouch anastomosis; APR: abdominoperineal resection; LRA: transanally double stapled low colorectal anastomosis; ECOG: Eastern Cooperative Oncology Group; SF-12: Short-Form 12; FACT-C: functional assessment of cancer therapy – colorectal; TNM: tumour, node, metastases; HADS: hospital anxiety and depression scale; PANSS: positive and negative syndrome scale.

**Table A5: Responsiveness results for bowel cancer, adapted from Longworth et al 2014. (34)**

| Author, year | Method of measuring responsiveness | Responsiveness results |
|---|---|---|
| Anderson and Palmer 1998(10) | Odds ratio for responses of EQ-5D dimensions between baseline and week 5 and 15 over the two groups<br><br>Odds ratio for responses of RSCL dimensions between baseline and week 5 and 15 over the two groups | Week 2 (changes from baseline):<br>RSCL scores (all dimensions and sub-divisions) significantly different in patients randomised to raltitrexed and patients randomised to 5-FU+LV arms with the exception of psychological symptoms and disease categories.<br>EQ-5D scores (4 dimensions and general health questions) highly significantly different in favour of raltitrexed .<br>Raltitrexed group three times less likely to have problems with mobility and usual activities than 5-FU+LV group (OR 2.9 and p<0.02). Ralaxitred group at least twice as likely to have a better general health (OR 2.3, p<0.001) and 2 to 3 times as capable of self-care as patients in the 5-FU+LV group but this result not significant.<br><br>Week 10:<br>No statistically significant differences between groups .<br>Non-significant trends in favour or reltitrexed on the EQ-5D scale and in total symptom advantages. |
| Doornebosch 2008(11) | Wilcoxon's signed rank test and Mann-Whitney U test for change scores within or between groups. Spearman rank correlations between change scores. | FISI mean scores showed significant post-op decrease. Greater decrease for patients with a tumour location within 7cm from the denatate line (p=0.01).<br>FIQL showed a significant post-op improvement in two of the four domains (embarrassment and lifestyle). The domains of lifestyle, coping and behaviour and embarrassment were correlated with the FISI. |
| Janson 2007(8) | Mean changes of scores between intervention groups (MANOVA analysis of change over time). | EQ-5D scores not significantly different between intervention groups.<br>EORTC QLQ-C30 scores showed significant benefit of LCR at the 2 and 4 week assessments. At the 12 week assessment, this was borderline significant. Significant benefit of LCR for role function found at two week assessment. |
| Sharma 2007(4) | Mean changes in EQ-5D scores before and after surgery.<br>Mean changes in HADS, PANAS, FACT-EW, EQ-VAS scores before and after surgery. | HADS depression score significantly higher in the 6 week post-discharge measure (3.6 vs 4.8, p<0.05)<br><br>HADS anxiety, PANAS positive and PANAS negative affect, FACT-C total, FACT-C trial outcome index and EQ-5D were not significantly different postoperatively. |
| Siena 2007(14) | EQ-5D scores by intervention group, FCSI scores by intervention group | FCSI change scores significant for both intervention and best supportive care groups.<br>EQ-5D only found significant change in intervention group. |

RSCL: Rotterdam symptom checklist; FU: fluorouracil; LV: leucovonin; EQ-5D: EuroQoL 5 dimensions; FISI: faecal incontinence severity index; FIQL: faecal incontinence quality of life; EORTC QLQ: European organisation for research and treatment of cancer core quality of life questionnaire; LCR: laparoscopic colon resection; HADS: hospital anxiety and depression scale; PANAS: positive and negative affect scale; FACT-C: functional assessment of cancer therapy – colorectal; FACT-EW: functional assessment of cancer therapy – emotional well-being; FCSI: FACT colorectal symptom index.

**Table A6: Mandatory fields collected in the bowel cancer NCA**

| |
|---|
| *PATIENT* |
| NHS number, Date Of birth, Postcode (at diagnosis), Gender |
| *TUMOUR* |
| Organisation first seen, Date of diagnosis, Source of referral, Major site ICD10, Performance status, Care plan intent, Planned cancer treatment type, No cancer treatment reason, T category, N category, M category |
| *SURGERY* |
| Provider organisation, ASA grade, Cancer treatment curability, Date of surgery, Surgical urgency mode of operation, Consultant, Primary procedure, Surgical access |
| *PATHOLOGY* |
| Status of circumferential excision margin, Number of nodes examined, Number of nodes positive, Final pathology T category, Final pathology N category, Final pathology radiology M category |
| *CHEMORADIOTHERAPY* |
| Pre-op initial provider organisation, Pre-op initial cancer treatment modality, Post-op provider organisation, Post-op cancer treatment modality |
| *PATIENT (Dataset refs: B1-B8)* |
| NHS number, Originating organisation code, Updating organisation code, Uploading organisation code, Batch ID, Batch record ID, Sex, Height, Weight, Post mortem (Y/N) |
| *TUMOUR (Dataset refs: B9-B48)* |
| NHS number, Care spell number, Originating organisation code, Updating organisation code, Uploading organisation code, Batch ID, Batch record ID, Place first seen organisation code, Date of clinical diagnosis, ICD10 major site code, Synchronous sites (caecum, appendix, ascending colon, hepatic flexure, transverse colon, splenic flexure, descending colon, sigmoid colon, recto/sigmoid, rectum), Height of tumour above anal verge, Modified Dukes' staging, Clinical intervention date – colonoscopy, Colonoscopy complications, Clinical intervention date - barium enema, Patient procedure result - barium enema, CT colonography, Clinical intervention date - CT scan, Patient procedure result - CT scan, Clinical intervention date - 1st MRI scan, Patient procedure result -1st MRI scan T stage, Patient procedure result -1st MRI Scan N stage, 1st MRI scan margin threatened – result, Clinical intervention date -2nd MRI Scan, Patient procedure result -2nd MRI scan T stage, Clinical intervention date - Endoanal ultrasound, Patient procedure result - endoanal ultrasound, Clinical intervention date - Abdominal ultrasound, Patient procedure result - abdominal ultrasound, Distant metastases: (liver, lung, bone, other) |
| *TREATMENT (Dataset refs: B49-B83)* |
| NHS number, Care spell number, Treatment ID, Originating organisation code, Updating organisation code, Uploading organisation code, Batch ID, Batch record ID, Surgery provider organisation code, Start date of 1st definitive procedure treatment, Reason no surgery performed, ASA grade, Thromboembolism prevention, Antibiotic infect prevention, Cancer treatment intent (curability), Complications of cancer, Anaesthetist grade, Patient procedure (anastomosis), Patient Procedure (stoma), Date stoma closed, Surgical access, Type of bowel division at laparoscopy, Type of anastomosis at laparoscopy, Morbidity code (Major postoperative complication, Morbidity code (Major laparoscopic specific complication), Early port site complication, Excision margin (positivity of cut colon or rectum margin), Distance of tumour to nearest cut bowel margin (mm), Excision margin (circumferential margins), Distance between cancer and circumferential margins (mm), Perforation or serosal involvement, Distance between lower end of tumour and resection margin in rectal and rectosigmoid tumours (mm), Distance between lower end of cancer and dentate line in APER specimens (mm), Site specific staging classification (pathological, Dukes' Staging), Teletherapy type given, Teletherapy trial, Chemotherapy trial |
| *FOLLOW UP (Dataset refs: B84-B96)* |
| NHS number, Care spell number, Follow up ID, Originating organisation code, Updating organisation code, Uploading organisation code, Batch ID, Batch record ID, Organisation code (follow up provider), Mode of follow-up, Primary tumour status (local recurrence), Local recurrence diagnosed by, Wound recurrence, Port site recurrence, Site of distance spread |

**Table A7: Optional fields collected in bowel cancer NCA** (WP1.3)

| |
|---|
| *PATIENT*[a] |
| Patient surname, Patient forename |
| *TUMOUR*[a] |
| Clinical nurse specialist indication, Synchronous cancer, Monitoring intent |
| *SURGERY*[a] |
| CPES anaerobic threshold, BMI, Immediate post-operative care |
| *PATHOLOGY*[a] |
| No additional fields |
| *CHEMORADIOTHERAPY*[a] |
| No additional fields |
| *PATIENT*[b] |
| Patient local identifier code, Patient forename, Patient surname, Postcode, Date of birth, Consultant code, Date of death, Cause of death |
| *TUMOUR*[b] |
| Date of diagnosis, Referral source, Diagnostic route, Date of referral receipt, Priority of referral to outpatients, Date of first hospital appointment, **Patient procedure results – colonoscopy**, Colonoscopy incomplete reason, Final pre-treatment T category, Final pre-treatment N category, Final pre-treatment M category, MDT discussion indicator |
| *TREATMENT*[b] |
| Colorectal nurse or stoma therapist seen, Date seen by colorectal nurse or stoma therapist, Procedure date (date of surgery), Theatre case start time (24hr), Surgical Urgency (mode of operation), Primary procedure name (OPCS), Code of responsible HCP (Surgeon GMC code), Grade of responsible HCP (grade of operating surgeon), Discharge date (hospital provider spell (Date of discharge or death), Organisation code (pathology provider), date specimen sample received, Investigation result date (date of report), Authorising pathologist GMC code, Service report status, Service report identifier, Synchronous cancer indicator, Invasive lesion site (cancer size, mm), Grade of differentiation, Histology (SNOMED), Nodes examined number (number of lymph nodes found), Nodes positive number (number of positive lymph nodes found), Cancer vascular or lymphatic invasion (extramural vascular invasion), T category (pathological), N category (pathological), M category (pathological), Site code of teletherapy treatment, Teletherapy consultant code, Start date teletherapy treatment course (radiotherapy start date), Sidte code (of cancer drug treatment (Hospital)), Consultant code, Drug treatment intent, Start date (anti-cancer drug regiment) |
| *FOLLOW UP*[b] |
| Clinical assessment date (cancer, date of follow-up), Metastatic status dist spread, **Treatment related morbidity** |

[a] collected via an excel spreadsheet; [b]Extracted from Pdf: National Bowel Cancer Audit Dataset v3.1 (19 October 2009), BMI: body mass index, MDT: multi-disciplinary team

**Reference List**


(1) Longworth L, Yang Y, Young T, Mulhern B, Hernandez AM, Mukuria C, et al. Use of generic and condition-specific measures of health-related quality of life in NICE decision-making: A systematic review, statistical modelling and survey. Health Technol Assess 201418(9):1-224. Available from: URL: http://www.journalslibrary.nihr.ac.uk/__data/assets/pdf_file/0008/108368/FullReport-hta18090.pdf

(2) Hadi M, Gibbons E, Fitzpatrick R. A structured review of patient-reported outcome measures (PROMs) for colorectal cancer. Report to the department of health, 2010. Patient-reported Outcome Measurement Group 2010.

(3) Shea BJ, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Methodol 2007;7:10.

(4) Sharma A, Sharp DM, Walker LG, Monson JR. Predictors of early postoperative quality of life after elective resection for colorectal cancer. Ann Surg Oncol 2007 Dec;14(12):3435-42.

(5) Doornebosch PG, Tollenaar RA, Gosselink MP, Stassen LP, Dijkhuis CM, Schouten WR, et al. Quality of life after transanal endoscopic microsurgery and total mesorectal excision in early rectal cancer. Colorectal Dis 2007 Jul;9(6):553-8.

(6) Gosselink MP, Busschbach JJ, Dijkhuis CM, Stassen LP, Hop WC, Schouten WR. Quality of life after total mesorectal excision for rectal cancer. Colorectal Dis 2006 Jan;8(1):15-22.

(7) Hamashima C. Long-term quality of life of postoperative rectal cancer patients. J Gastroenterol Hepatol 2002 May;17(5):571-6.

(8) Janson M, Lindholm E, Anderberg B, Haglind E. Randomized trial of health-related quality of life after open and laparoscopic surgery for colon cancer. Surg Endosc 2007 May;21(5):747-53.

(9) Wilson TR, Alexander DJ, Kind P. Measurement of health-related quality of life in the early follow-up of colon and rectal cancer. Dis Colon Rectum 2006 Nov;49(11):1692-702.

(10) Anderson H, Palmer MK. Measuring quality of life: impact of chemotherapy for advanced colorectal cancer. Experience from two recent large phase III trials. Br J Cancer 1998;77 Suppl 2:9-14.

(11) Doornebosch PG, Gosselink MP, Neijenhuis PA, Schouten WR, Tollenaar RA, de Graaf EJ. Impact of transanal endoscopic microsurgery on functional outcome and quality of life. Int J Colorectal Dis 2008 Jul;23(7):709-13.

(12) Ramsey SD, Andersen MR, Etzioni R, Moinpour C, Peacock S, Potosky A, et al. Quality of life in survivors of colorectal carcinoma. Cancer 2000 Mar 15;88(6):1294-303.

(13) Colwell HH, Mathias SD, Turner MP, Lu J, Wright N, Peeters M, et al. Psychometric evaluation of the FACT Colorectal Cancer Symptom Index (FCSI-9): reliability, validity, responsiveness, and clinical meaningfulness. Oncologist 2010;15(3):308-16.

(14) Siena S, Peeters M, Van CE, Humblet Y, Conte P, Bajetta E, et al. Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. Br J Cancer 2007 Dec 3;97(11):1469-74.

(15) Aaronson NK, Ahmedzai S, Bergman B, Bullinger M, Cull A, Duez NJ, et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. J Natl Cancer Inst 1993 Mar 3;85(5):365-76.

(16) Rowen D, Brazier J, Young T, Gaugris S, Craig BM, King MT, et al. Deriving a preference-based measure for cancer using the EORTC QLQ-C30. Value Health 2011 Jul;14(5):721-31.

(17) Sprangers MA, te VA, Aaronson NK. The construction and testing of the EORTC colorectal cancer-specific quality of life questionnaire module (QLQ-CR38). European Organization for Research and Treatment of Cancer Study Group on Quality of Life. Eur J Cancer 1999 Feb;35(2):238-47.

(18) Whistance RN, Conroy T, Chie W, Costantini A, Sezer O, Koller M, et al. Clinical and psychometric validation of the EORTC QLQ-CR29 questionnaire module to assess health-related quality of life in patients with colorectal cancer. Eur J Cancer 2009 Nov;45(17):3017-26.

(19) Gujral S, Conroy T, Fleissner C, Sezer O, King PM, Avery KN, et al. Assessing quality of life in patients with colorectal cancer: an update of the EORTC quality of life questionnaire. Eur J Cancer 2007 Jul;43(10):1564-73.

(20) Arraras JI, Manterola A, Hernandez B, Arias dl, V, Martinez M, Vila M, et al. The EORTC information questionnaire, EORTC QLQ-INFO25. Validation study for Spanish patients. Clin Transl Oncol 2011 Jun;13(6):401-10.

(21) Khazaeli N, Golshiri P, Farajzadegan Z, Hemati S, Amouheidari A, Hakimian MR, et al. Evaluating the validity and reliability of Persian version of the European organization for research and treatment of cancer quality of life questionnaire for colorectal cancer (EORTC QLQ-CR29). Journal of Isfahan Medical School 2014;32(276).

(22) Roche. Acheiving Clinical Excellence in the Treatment of Metastatic Colorectal Cancer TA118. 2014. Report No.: TA118.

(23) Sanofi-Synthelabo Ltd. The use of oxaliplatin for the treatment of advanced colorectal cancer (review of NICE guidance #33) TA93. 2014.

(24) Ward S, Kaltenthaler E, Cowan J, Brewer N. A review of the evidence for the clinical and cost-effectiveness of Capecitabine and Tegafur with Uracil for the treatment of metastatic colorectal cancer TA61. 2002.

(25) Lloyd-Jones M, Hummel S, Bansback N. A review of the evidence for the clinical and cost-effectiveness of Irinotecan, Oxaliplatin and Raltitrexed for the treatment of advanced colorectal cancer TA33. 2001.

(26) Hoyle M, Crathorne L, Peters J, Jones-Hughes T, Cooper C, Napier M, et al. The effectiveness and cost-effectiveness of cetuximab (mono- or combination chemotherapy), bevacizumab (combination with (non-oxaliplatin chemotherapy) and panitumumab (monotherapy) for the treatment of metastatic colorectal cancer after first-

line chemotherapy (review of technology appraisal 150 and part review of technology appraisal 118): a systematic review and economic model TA242. 2011.

(27) Murray A, Lourenco T, de Verteuil R, Hernandez R, Fraser C, McKinley A, et al. Systematic review of the clinical effectivenesss and cost-effectiveness of laparascopic surgery for colorectal cancer TA105. 2005.

(28) Pandor A, Eggington S, Paisley S, Tappenden P, Sutcliffe P. The use of oxaliplatin and capecitabine for the aduvant treatment of colon cancer. National Institute for Health and Clinical Excellence; 2005.

(29) Whyte S, Pandor A, Stevenson M, Rees A. Bevacizumab in combination with Fluoropyrimidine-based chemotherapy for the first line treatment of metastatic colorectal cancer TA212. 2009.

(30) Merck Sorono. Erbitux (Cetuximab) for the first line treatment of metastatic colorectal cancer TA176. 2008.

(31) Sanofi UK.
Aflibercept (Zaltrap®) for the treatment of metastatic colorectal cancer (mCRC) previously treated with an oxaliplatin containing regimen ID514. 2013.

(32) National Clinical Guideline Centre, National. Colorectal Cancer: the diagnosis and management of colorectal cancer CG131. 2011.

(33) Eggington S, Pandor A, Paisley S, Tappenden P, Sutcliffe P. The use of Oxaliplatin and Capecitabine for the adjuvant treatment of colon cancer TA100. 2005.

(34) Longworth L, Singh J, Brazier J. An evaluation of the performance of EQ-5D: a review of reviews of psychometric properties. Unpublished report to EuroQol Group, editor. 2014.