

Patient-Reported

**PATIENT-
REPORTED
OUTCOME
MEASUREMENT
GROUP, OXFORD**

**A STRUCTURED REVIEW
OF
PATIENT-REPORTED
OUTCOME MEASURES
(PROMs)
FOR COLORECTAL
CANCER**

**Report to the Department of
Health, 2010**



O

U

T

C

O

M

E

Measures

**A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME
MEASURES FOR COLORECTAL CANCER**

Monica Hadi
Elizabeth Gibbons
Ray Fitzpatrick

*Patient-reported Outcome Measurement Group
Department of Public Health
University of Oxford*

April 2010

Copies of this report can be obtained from:

Elizabeth Gibbons
Senior Research Officer
Patient-reported Outcome Measurement Group

tel: +44 (0)1865 289402
e-mail: elizabeth.gibbons@dphpc.ox.ac.uk

A STRUCTURED REVIEW OF PATIENT-REPORTED OUTCOME MEASURES FOR COLORECTAL CANCER, 2010

CONTENTS

Executive summary	1
Introduction, methods, search terms and results	2
Generic PROMs	7
Preference-based PROMs	9
Cancer-specific PROMs	10
Other PROMs in Colorectal Cancer	15
Summary of evidence and recommendations	17
Appendix A – Sources for search and search strategy	21
Appendix B – Appraisal of the methodological quality of PROMs	23
Appendix C – Generic and Utility PROMs (description)	25
Appendix D – Cancer Specific PROMs (description)	28
Appendix E – Licensing and contact details	32
References	33

EXECUTIVE SUMMARY

Aims of the report

The aims of this report are to review the evidence of Patient-reported Outcome Measure (PROMs) for people with colorectal cancer and provide a short-list of the most promising generic and cancer-specific instruments. The methods of the review are described and the results of the search, including sources and search terms used to identify specific published research. Details of this evidence are presented firstly for generic PROMs evaluated with people with colorectal cancer, followed by condition-specific PROMs. The report concludes with discussion and recommendations.

Results

Two generic instruments, which have been evaluated with colorectal cancer, were identified in this review:

1. Medical Outcomes Study 36-Item Health Survey (SF-36)
2. Medical Outcomes Study 12-Item Health Survey (SF-12)

One Preference based measure, which has been evaluated with colorectal cancer, was identified in this review:

1. European Quality of Life Questionnaire (EuroQol; EQ-5D)

One general cancer-specific instrument was identified in this review:

1. European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30).

Two colorectal cancer-specific instruments were identified:

2. European Organization for Research and Treatment of Cancer Colorectal Cancer Specific Quality of Life Questionnaire (EORTC QLQ-CR38)
3. Functional Assessment of Cancer Therapy – Colorectal (FACT-C)

Recommendations

Based on the volume of evaluations and good measurement and operational characteristics, the following are highlighted as promising PROMs for potential piloting in the NHS in people with colorectal cancer:

Generic:

SF-12

Preference-based:

EQ-5D

Cancer-specific:

EORTC QLQ-C30

FACT-C

As survival rates in patients with colorectal cancer have improved, attention may need to be given to longer term issues of survivorship.

INTRODUCTION

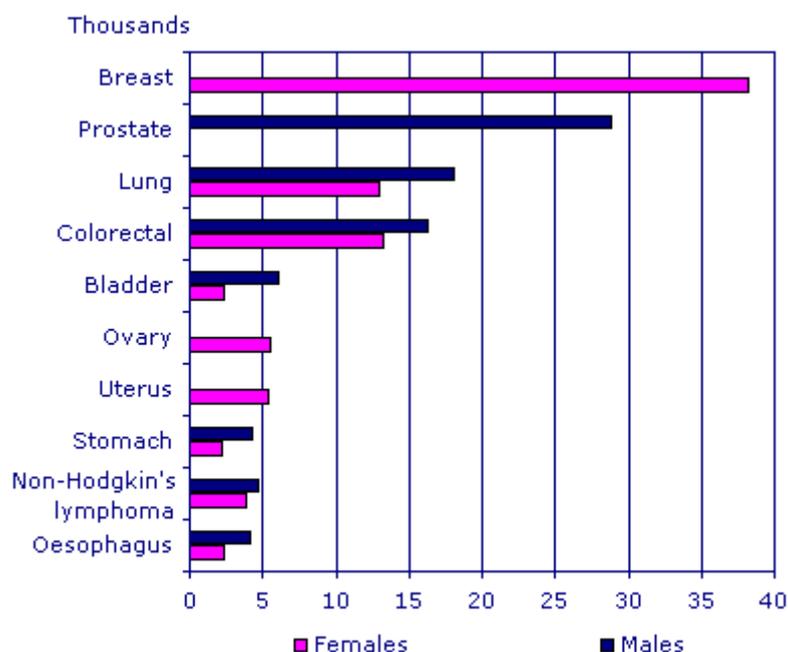
Background

Patient-reported outcome measures (PROMs) offer enormous potential to improve the quality and results of health services. They provide validated evidence of health from the point of view of the user or patient. They may be used to assess levels of health and need in populations, and in users of services, and over time they can provide evidence of the outcomes of services for the purposes of audit, quality assurance and comparative performance evaluation. They may also improve the quality of interactions between health professionals and individual service users.

Lord Darzi's Interim Report on the future of the NHS recommends that patient-reported outcome measures (PROMs) should have a greater role in the NHS (Darzi, 2007). The new Standard NHS Contract for Acute Services, introduced in April 2008, included a requirement to report from April 2009 on patient-reported outcome measures (PROMs) for patients undergoing Primary Unilateral Hip or Knee replacements, Groin Hernia surgery or Varicose Vein. Furthermore, Lord Darzi's report 'High Quality Care for All' (2008) outlines policy regarding payments to hospitals based on quality measures as well as volume. These measures include PROMs as a reflection of patients' experiences and views. Guidance has now been issued regarding the routine collection of PROMs for the selected elective procedures (Department of Health, 2008) and since April 2009, the collection of PROMs for the selected elective procedures has been implemented and is ongoing.

In light of recent policy to include PROMs as an important quality indicator, the Department of Health now seeks guidance on PROMs which can be applied in patients with cancer and commissioned the Patient-reported Outcome Measurement Group, Oxford, to review the evidence of PROMs for selected cancers. It is proposed that the most common cancers, as identified via the Office for National Statistics, should be the subject of review in terms of most promising PROMS. Breast, lung, colorectal and prostate cancer are highlighted as being the four most common cancers, accounting for half of the 239,000 new cases of malignant cancer (excluding non-melanoma skin cancer) registered in England in 2005 (Figure 1). On scrutinising cumulative incidence data from the cancer registry of the Oxford region, findings support that these four cancers are the most common. According to the Department of Health's Cancer Reform Strategy (2007), which aims to place the patient at the centre of cancer services, a 'vision 2012' has been created for each of these four cancer types, highlighting the progress that it is hoped will be made by 2012 in terms of the cancer pathway. Underlying these visions are the aims to achieve full implementation of improving outcomes guidance. In this context, PROMS are an important resource to monitor cancer outcomes.

Figure 1: Incidence of the major cancers, 2005, England (ONS, 2007)



Colorectal Cancer

Colorectal cancer, also known as colon/rectal/large bowel cancer, includes cancerous growths in the colon, rectum and appendix. It is one of the four most common cancers in the UK; each year, over 30,000 new cases are diagnosed¹. Colorectal cancers arise from adenomatous polyps in the colon. These are usually benign, but some can develop into cancer over time. Localized colon cancer is usually diagnosed through colonoscopy. Invasive cancers that are confined within the wall of the colon (TNM stages I and II) are curable with surgery. If untreated, they spread to regional lymph nodes (stage III), where up to 73% are curable by surgery and chemotherapy. Cancer that metastasizes to distant sites (stage IV) is usually not curable, although chemotherapy can extend survival, and in rare cases, surgery and chemotherapy together have seen patients through to a cure. Survival has doubled over the last 30 years as a result of early diagnosis (NICE, 2004).

Colorectal cancer is thus a priority on the government health agenda and has been for some time. The national institute of Health and Clinical Excellence (NICE, 2004) have published 'Improving outcomes in colorectal cancer'; a guidance on cancer services providing key recommendations for the improvement of health outcomes in colorectal cancer. Recommendations have generally focused on action to improve recognition of symptoms in primary care and community and to improve access to screening. More importantly, the role of prompt referral and management of patients with rectal cancer have also been highlighted. Better information via patient-reported outcomes could enhance quality of care, inform commissioning, and promote patient choice. The following review provides current information available on PROMs used with colorectal cancer patients.

¹ Source: data on Office for National Statistics website < www.nationalstatistics.gov.uk >

METHODS AND SEARCH STRATEGY

Aim of the report

The aim of this report is to identify Patient-reported Outcome Measures (PROMs) which have been evaluated with patients with colorectal cancer.

Structure of the report

The methods of the review are described, including search strategies and search terms used to identify relevant published research regarding PROMs for people with colorectal cancer. Details of this evidence are presented for generic PROMs, Utility measures and cancer-specific instruments. The report concludes with discussion and recommendations.

Methods

Methods adopted were as described in previous reviews performed by the PROM group, Oxford. Comprehensive searches were conducted; articles retrieved were assessed for relevance and evidence of measurement performance and operational characteristics abstracted for each PROM identified.

a) Search terms and results: identification of articles

The searches were conducted using three main sources.

Records in the bibliography database were searched up to December 2005. This database was compiled by the PROM group with funding from the Department of Health and the Information Centre, and hosted by the University of Oxford¹.

The Ovid search engine was used to explore a number of relevant databases² from January 2006 until February 2010, using a comprehensive search strategy (see appendix A).

Hand searching of titles of key journals from October 2009 was conducted. The following journals were selected:

- Health and Quality of Life Outcomes
- Quality of Life Research.
- Journal of Clinical Oncology
- British Journal of Cancer
- Cancer
- Disease of Colon and Rectum

² Four databases were searched using Ovid: AMED (Allied and Complementary Medicine), EMBASE, PsycInfo, Ovid MEDLINE(R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE(R).

The following supplementary sources were searched:

- The National Institute for Health Research: Health Technology Assessment Programme
- The Cochrane library <http://www.thecochranelibrary.com/>
- The EQ-5D website; reference search facility (<http://www.euroqol.org/>)

The number of relevant articles identified through each source is shown in Table 1.

b) Inclusion criteria

Published articles were included if they provided evidence of measurement and/or practical properties of relevant PROMs (Fitzpatrick et al., 1998).

Population

- Patients with colorectal cancer
- English-speaking populations.

Study design selection

- studies where a principal PROM is being evaluated;
- studies evaluating several PROMs concurrently;
- trials or studies evaluating the effectiveness of interventions; where a PROM is used as an endpoint;
- prospective studies measuring patient-reported outcomes where data is available for a PROM in terms of measurement performance or operational characteristics.

Specific inclusion criteria for generic, preference-based and condition-specific instruments

- the instrument is patient-reported;
- there is published evidence of measurement reliability, validity or responsiveness following completion in the specified patient population;
- evidence is available from English-language publications, and instrument evaluations conducted in populations within the UK, North America, or Australasia;
- the instrument will ideally be multi-dimensional. It is at the reviewer's discretion to include RPOMs which are specific to a health condition but have a narrow focus, for example, a specific dimension of health, such as symptoms.

Exclusions

- studies using clinician-rated instruments;
- studies evaluating the performance of non-patient reported measures of functioning or health status where a PROM is used as a comparator;
- studies with very small samples, i.e. fewer than 40 participants
- studies using incomplete versions of instruments.

c) Data extraction

For all PROMs included in the review, evidence is reported for the following measurement criteria:

- reliability
- validity
- responsiveness
- precision

Operational characteristics, such as patient acceptability and feasibility of administration for staff, are also reported.

d) Assessment of methodological quality of PROMs

Assessment and evaluation of the PROMs was performed by means of the criteria described in Appendix B

Results

Searches identified nearly 1330 potentially relevant records. When assessed against the review inclusion criteria, 35 articles were included in the review (Table 1).

Table 1: Number of articles identified by the literature review

<i>Source</i>	<i>Results of search</i>	<i>Number of articles included in review</i>
PROM bibliography: 30,350	188	6
OVID 2006-February 2010	1139	28
Hand searching	3	1
TOTAL	1330	35

RESULTS: Generic PROMs evaluated in colorectal cancer

Two generic instruments were identified:

- a) SF-36
- b) SF-12

a. SF-36

The Medical Outcomes Study (MOS) Short Form 36-item Health Survey (SF-36) is intended for application in a wide range of conditions and with the general population. The instrument assess health across eight domains (Ware., 1997), namely Bodily pain (BP), General Health Perceptions (GH), Mental health (MH), Physical functioning (PF), Role limitations due to emotional health problems (RE), Role limitations due to physical health problems (RP), Social functioning (SF), and Vitality (VT). An additional health transition item, not included in the final score, assesses change in health.

A total of 4 articles evaluating the SF-36 are included in the review. None of these were conducted in the UK.

Validity

Discriminant validity was demonstrated by Anthony et al. (2003) as patients experiencing surgical complications had significantly lower scores on the SF, GH and MH domains as well as on the Mental Health Component Summary Score (MCS). Older age has been shown to be associated with worse VT scores (Mastracci et al., 2006) and PF scores (Kariv et al., 2008; Reeve et al., 2009)

In a longitudinal study, Reeve et al. (2009) found that cancer patients reported a significantly greater decline in both the Mental and Physical component summary scores when compared to matched control subjects. Furthermore, Kariv et al. (2008) found that the PF domain of the SF-36 to discriminate between patients who received post-operative radiotherapy compared with those who did not receive radiotherapy.

Responsiveness

In a 2 year longitudinal investigating of HRQOL pre- and post cancer diagnosis, Reeve et al. (2009) observed a statistically significant decline in the RP domain.

b. SF-12

In response to the need to produce a shorter instrument that could be completed more rapidly, the developers of the Medical Outcomes Study (MOS) 36-item Short Form Health Survey (SF-36) produced the 12-item Short Form Health Survey (SF-12) (Ware et al., 1995). The same eight domains as the SF-36 are assessed and categorical response scales are used.

A total of 4 articles evaluating the SF-12 are included in the review. Of these, 2 were conducted in the UK (Wilson et al., 2006; Gall et al., 2007).

Validity

There is some support for the discriminant ability of the SF-12 domains. In a UK study by Wilson et al. (2006), the MCS domain was shown to discriminate between groups of patients based on age whilst the PCS domain was able to discriminate by age, cancer site (colon vs. rectal) and health care utilisation (length of stay in hospital). In addition, Gall et al. (2007) reported older age to be related to significantly lower PCS scores providing further support of discriminative validity.

Wilson et al. (2006) reported that the SF-12 in general performed slightly better than condition specific instruments (QLQ-C30 and FACT-C) in determining HrQoL differences based on general health-related constructs, such as performance status. In contrast, the condition specific instruments outperformed the SF-12 in detecting larger differences in HrQoL between groups based on disease-specific symptoms, such as diarrhoea and constipation.

Responsiveness

In a UK longitudinal study, Wilson et al. (2008) reported improved postoperative scores in the MH, PF, RP and SF domains at 3-6 months follow-up.

Acceptability

Wilson et al. (2008) reported a postal response rate of 88.6% at 12 months, 64.8% at 18 months and 40% at 24 months. Wilson et al (2006) reported limited missing data at 3.7%.

RESULTS: Preference-based PROMs evaluated in colorectal cancer

One preference-based instrument was identified:

a. EQ-5D

The European Quality of Life instrument (EuroQol), now generally known as the EQ-5D, is divided into two sections: the EQ-5D index and the EQ thermometer (EQ-5D VAS). The EQ-5D assesses health across five domains: anxiety/depression (AD), mobility (M), pain/discomfort (PD), self-care (SC), and usual activities (UA). The EQ thermometer is a single 20cm vertical visual analogue scales with a range of 0 to 100, where 0 is the worst and 100 the best imaginable health.

A total of 3 articles evaluating the EQ-5D are included in the review. Of these, 2 were conducted in the UK (Wilson et al., 2006; Sharma et al., 2007)

Validity

Discriminant validity was demonstrated by Best et al. (2010) as cancer patients reported significantly higher scores on the AD domain when compared to a community sample. In a UK study by Wilson et al. (2006), the EQ-5D was shown to discriminate between groups of patients based on cancer site (colon vs. rectal) and health care utilisation whilst the EQ thermometer was able to discriminate between groups of patients by age and symptom (constipation).

In a UK study, Wilson et al. (2006) reported that the EQ-5D in general performed slightly better than the condition specific instruments (QLQ-C30 and FACT-C) in determining HrQoL differences based on general health-related constructs, such as performance status. In contrast, the condition specific instruments outperformed the EQ-5D in detecting larger differences in HrQoL between groups based on disease-specific symptoms, such as diarrhoea and constipation.

Sharma et al. (2007) demonstrated predictive validity as a significant correlation was found between pre- and postoperative EQ-5D VAS scores.

Acceptability

Wilson et al. (2006) reported limited missing data at 4% for EQ-5D index and 3.5% for EQ-5D VAS.

RESULTS: Cancer-specific PROMs in colorectal cancer

One general cancer-specific instrument was identified:

a) EORTC QLQ-C30

Two colorectal cancer-specific instruments were identified:

b) EORTC QLQ-CR38

c) FACT-C

a. European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30)

The EORTC QLQ-C30 (Aronson et al., 1993) is a 30-item cancer-specific instrument. Multi-trait scaling was used to create five functional domain scales: Physical (P), Role (R), Emotional (E), Social (S) and Cognitive (C); two items evaluate global health status (GHS); in addition three Symptom scales assess: Fatigue, Pain and Emesis; and six single items assess further symptoms.

A total of 12 articles evaluating the EORTC QLQ-C30 are included in the review. Half of these were conducted in the UK (Wilson et al., 2006; King et al., 2006, 2008; Knowles et al., 2007; Jayne et al., 2007; Yau et al., 2009).

Validity

The functional domains of the EORTC QLQ-C30 have been shown to discriminate between known groups based on age (Wilson et al., 2006; Mastracci et al., 2006), health care utilisation (length of stay in hospital) and symptom (constipation) (Wilson et al., 2006). In addition, Neuman et al. (2007) reported that, when compared with the published population norms, colorectal cancer patients reported more Symptoms of constipation, diarrhoea and experienced financial difficulty reflected in the symptom subscales.

Furthermore, evidence from several randomised controlled trials have demonstrated discriminant validity. The P and GHS domains have been shown to discriminate between groups based on type of clinical intervention (Jonker et al., 2006; Au et al., 2009) and malnutrition (Gupta et al., 2006).

Similarly, the S domain and one of the symptom scales (constipation) have been shown to discriminate between groups based on type of surgery (laparoscopic or conventional) (Jayne et al., 2007), the presence of stoma and type of stoma (temporary or permanent) (Yau et al., 2009). The P domain has also been shown to discriminate between genders (Neuman et al., 2007).

A UK randomised trial by Yau et al. (2009) reported no differences in scores between stoma and non-stoma patients on the GHS, social and role functioning at baseline, however; a significant difference in scores was noted between the two groups during chemotherapy as stoma patients scored lower on social and role functioning domains as well as symptoms of constipation.

In contrast, a UK trial by King et al. (2006; 2008) reported no significant differences between the groups (laparoscopic or open surgery) in any of the QoL components of the EORTC QLQ-C30 post-surgery. However, as both groups received some type of surgery, the questionnaire may be indicating improved HrQoL in both study arms.

In a UK study, Wilson et al. (2006) reported that the QLQ-C30 outperformed generic measures (SF-12 and EQ-5D) in detecting larger differences in HrQoL between groups based on disease-specific symptoms, such as diarrhoea and constipation. In contrast, the generic measures performed slightly better than the QLQ-C30 in determining HrQoL differences based on general health-related constructs, such as performance status.

Responsiveness

Responsiveness evidence is reported from several randomised controlled trials. In a study to assess the feasibility of a follow-up programme led by nurse specialists, Knowles et al. (2007) reported that QLQ-C30 Physical, Social, Role and GHS domains along with symptoms of fatigue, pain, sleep disturbance, appetite loss and diarrhoea improved at 12 months follow-up. Change in scores was detected in a UK randomised trial by Yau et al. (2009) which reported that both patients with and without stoma had improved scores on all functioning domains at one and three years after chemotherapy.

In addition, in surgical trials, Jayne et al. (2007) reported that the P, E, S and C domains were responsive to change at 2 weeks follow-up whilst King et al. (2006) reported that the R and P domains along with symptoms pain and fatigue to be responsiveness to change 2 weeks post-surgery, returning to baseline levels at 6 weeks.

Acceptability

Wilson et al. (2006) reported negligible missing data (0.5%). Au et al. (2009) reported compliance at baseline to be 94%, declining to 47-67% at 16 weeks follow-up depending on study arm. Similarly figures were shown by Jonker et al. (2006) with a 94% response rate at baseline dropping to 61-82% at 8 weeks and 43-67% at 16 weeks depending on intervention group. Neuman et al. (2007) reported that 67% of patients completed the questionnaire at baseline and that missing data was less than 1% (1 of 124 patients).

A high response rate of 80% has been reported at 2-6 weeks (King et al., 2006) and 65 months postal follow-up (Hassan et al., 2006). Jayne et al. (2007) indicated that 347 (69%) of 500 patients completed the questionnaire at 3 years postoperatively. Furthermore, Yau et al. (2009) reported a compliance rate of 72% at baseline, 66% during chemotherapy and later 78% and 85% at one and three years follow-up, respectively.

b. European Organization for Research and Treatment of Cancer Colorectal Cancer Specific Quality of Life Questionnaire (EORTC QLQ-CR38)

The EORTC QLQ-C38 (Sprangers et al., 1999) is a 38-item colorectal cancer-specific module to be administered in addition to the EORTC QLQ-C30 questionnaire. The 38 items are added to the 30 core items of the QLQ-C30, becoming a 68 item questionnaire. Including the domains of the QLQ-C30, the QLQ-CR38 adds two functional scales (body image (BI) and sexual function (SF)), seven symptom scales (micturition problems (MP), gastrointestinal tract symptoms (GT), chemotherapy side effects (GS), defecation problems (DP), stoma-related problems (SP), and male and female sexual problems (MS & FS)), and three single-item measures (sexual enjoyment, weight loss, and future perspective).

A total of 7 articles evaluating the EORTC QLQ-CR38 are included in the review. Of these, 4 were conducted in the UK (King et al., 2006, 2008; Knowles et al., 2007; Jayne et al., 2007).

Validity

The GT and DP subscales have been reported to discriminate between patients grouped by age (Mastracci et al., 2006; Neuman et al., 2007). In addition, the GT subscale also been reported to discriminate based on gender whilst the DP discriminated based on type of surgical procedure and whether radiotherapy was received (Neuman et al., 2007).

In a randomised controlled trial of patients receiving pelvic radiotherapy (either preoperative, postoperative or no pelvic radiotherapy), Hassan et al. (2006) reported that the FP, MP, CS and DF scales of the EORTC QLQ-CR38 were able to detect a significant difference between groups at 65 months follow-up (median). Similarly, Jayne et al. (2007) reported that the MS scales was able to discriminate between groups based on types of surgery (laparoscopic versus conventional).

In contrast, a UK trial by King et al. (2006; 2008) no significant differences were reported between the groups (laparoscopic or open surgery) in any of the QoL components of the EORTC QLQ-CR38 post-surgery. However, as both groups received some type of surgery, the questionnaire may be indicating improved HrQoL in both study arms.

Responsiveness

Jayne et al. (2007) reported that BI and SF scales were responsive to change at 2 weeks follow-up. Knowles et al. (2007) reported significantly improved scores on the QLQ-CR-38 SF, DP at 12 months follow-up.

Acceptability

Neuman et al. (2007) reported that 67% of patients completed the questionnaire at baseline and that missing data was less than 1% (1 of 124 patients). King et al. (2008) reported that compliance rates for the EORTC QLQ-CR38 were over 95% at each time point (up to 1 year follow-up) and similarly, Jayne et al. (2007) reported that 347 (69%) of 500 patients completed the questionnaire at 3 years postoperatively.

The EORTC QLQ-CR38 has been widely used in many trials and research settings; however, its application alongside the QLQ-C30 has been inconsistent across studies. The reader should note that the authors do not advocate scoring by domain(s) or using only one domain of the instrument; the QLQ-C30 and the QLQ-CR38 are intended to be used and scored as one single instrument. Conversely, all studies that have been included in this review for the QLQ-CR38 have utilised the two instruments separately and analysed them as separate instruments. Therefore, results for the QLQ-CR38 alone should be interpreted with caution.

Furthermore, the QLQ-CR38 has never been formally evaluated. Work to revise the instrument began a few years ago and a shorter version, QLQ-CR29 was developed and preliminarily evaluated in an international multicentre trial with the aim to utilise the questionnaire in clinical trials with colorectal cancer patients (Whistance et al., 2009). There was strong support for its psychometric properties and subsequently, the instrument is now recommended by the developers as a valid and reliable PROM to be used with colorectal cancer patients. However, no published UK or English language evidence is reported yet.

c. FACT-C

The FACT-C (Ward et al., 1999) is a 37 item colorectal cancer-specific scale that supplements the general version; FACT-G (Cella et al., 1993). The instrument consists of 5 subscales: physical wellbeing (PW), social and family wellbeing (SW), emotional wellbeing (EW), functional wellbeing (FW) and the Colorectal Cancer Subscale (CCS). Scores can be produced as a combined total of all domains (FACT-C total), the Colorectal Cancer Score (CCS) and a Treatment Outcome Index (TOI) can be calculated by summing the FACT-G physical and functional domains and the CCS. Patients rate each item on the questionnaire from 0 to 4, where a higher note denotes a better quality of life.

A total of 11 articles evaluating the FACT-C are included in the review. Of these, 5 were conducted in the UK (Anthony et al., 2003; Twelves et al., 2006; Wilson et al., 2006; Sharma et al., 2007; Simon et al., 2009).

Reliability

Internal consistency reliability was reported by Ward et al. (1999) for all the FACT-subscales as measured from 2 separate samples (n= 60 and n=63). High internal consistency (Cronbach's alpha > 0.7) have been reported for PW, SW, EW, FW, CCS domains as well as for the FACT-C total and TOI. In addition, higher internal consistency was reported for FACT-C (0.91) than for FACT-G (0.88).

Similarly, Yost et al. (2005) reported internal consistency separately for three different samples (observational, preliminary and clinical trial sample). Overall, Cronbach's alpha was reported between 0.87-0.91 for TOI, 0.87-0.92 for the FACT-C total and a low 0.59-0.76 for CCS

Validity

Evidence of convergent validity has been demonstrated with significant positive correlations between FACT-C and its subscales and the FLIC (0.74) where higher scores indicate better QoL (Ward et al., 1999).

The Fact-C scores discriminated between groups of patients based on age, cancer site (colon vs. rectal), health care utilisation (length of stay in hospital) and symptom (constipation) (Wilson et al., 2006), stage of disease and gender (Simon et al., 2009), and surgical complications (Anthony et al., 2003).

Similarly, Sharma et al. (2007) indicated that patients who had stoma reported significantly lower scores on the EW and the CCS domains. Furthermore, the duration of postoperative stay showed significant negative correlation with FW and CCS scores whilst the presence of postoperative morbidity was found to have significant association with FW scores.

In contrast, some intervention studies have found no significant differences between therapy groups at follow-up (Twelves et al., 2006; Kabbinavar et al., 2008; Stephenson et al., 2009). However; it can be argued that the insignificant results from these studies may reflect the lack of significant difference in scores between type of intervention rather than the FACT-C's discriminative ability or the ability to detect change.

More recently, Wilson et al. (2006) reported that the FACT-C outperformed generic measures (SF-12 and EQ-5D) in detecting larger differences in HrQoL between groups based on disease-specific symptoms, such as diarrhoea and constipation. In contrast, the generic measures performed slightly better than the FACT-C in determining HrQoL differences based on general health-related constructs, such as performance status.

Sharma et al. (2007) demonstrated predictive validity as significant correlations were found between pre- and postoperative PW, SW, EW, FW and CCS domain scores.

In addition, Steginga et al. (2009) reported that QoL 6 months post-diagnosis (as defined by FACT-C scores) was the strongest predictor of QoL at 24 months post-diagnosis, for overall QoL and all FACT-C domain-specific subscales.

Responsiveness

Responsiveness to change at 2 months was initially reported by the developers in a sample of 40 patients (Ward et al., 1999). Minimally important differences (MID) have been calculated for FACT-C in Yost et al. (2005) and confirmed using three different statistical analyses (distribution based methods) and anchor based methods. MID scores ranged from 1-2 points for the CCS, 4-6 points for the TOI and 5-8 points for the FACT-C total. The preliminary and confirmatory MIDs were in close agreement, however; MID ranges for the observational sample tended to be slightly lower than those for the preliminary and clinical trial samples.

Acceptability

Steinga et al. (2009) reported a postal response rate of 53.2% at baseline whilst Stephenson et al. (2009) reported a 62% postal response rate at six months. Wilson et al. (2006) reported minimal missing data (7%).

Results: Other PROMs in colorectal cancer

Several additional PROMs were identified in the search strategy, but there was insufficient evidence to highlight any in the review. A short summary of these instruments is provided below.

Symptoms in Colorectal cancer

Patients with cancer often experience high symptom burden. This is reflected in the number of symptom measures available. Twenty-one cancer symptom instruments were identified in a systematic review (Kirkova et al., 2006). Of these, 18 were self-reported and most contained less than 20 items. None were specific to colorectal cancer nor was there any substantive evidence reported for any of these measures in relation to people with colorectal cancer.

Several other measures are available to assess the prevalence and related distress of symptoms. These tend to be either general symptom measures which have been developed with a population of people with chronic illness including cancer and others specific to cancer and cancer sites. Examples include the following:

- Edmonton Symptom Assessment Scale (general symptoms) (Bruera et al., 1991)
- Symptom Distress Scale (cancer related symptoms) (McCorkle, 1987).
- M.D. Anderson Symptom Distress Inventory (cancer related symptoms) (Cleeland et al., 2000).

This review reports some limited evidence for one general symptom measure: The Memorial Symptom Assessment Scale (MSAS) and the Condensed MSAS. These have been evaluated with people with bowel cancer with some limited evidence of reliability and validity. However, it is very narrow in its focus of generic symptoms specific to patients with cancer, assessing symptom prevalence and distress and has limited value for the evaluation of services (Portneoy et al., 1994; Chang et al., 2004).

Three other measures were identified in this review with limited evidence

Ferran's Quality of Life Index questionnaire measures overall QoL and QoL in four domains: Health and physical, Social and economic, Psychological and spiritual, and Family. Responses are obtained on a 6 point Likert scale in 2 parts: satisfaction with aspects of life and the importance of these aspects. Scores range from 0 to 30 with higher scores indicating a better QoL. Scores are also determined by weighted satisfaction responses with importance responses. Several evaluations with patients with cancer are reported; some with patients with colorectal cancer specifically in terms of predictive validity. The Health and physical sub-scale scores have been found to be significantly predictive of survival (Lis et al., 2006). Statistically significant differences in survival have been reported for people with different cancers including colorectal using score cut-off values. For example, the Health and function scores ≤ 17.4 and ≥ 17.4 were predictive of survival with the median survival being 9.5 and 23 months respectively ($p < 0.001$) (Lis et al., 2007).

The Functional Living Index-Cancer (FLIC) is a 22 item instrument with sub-scales relating to physical well-being and ability, psychological well-being, hardship due to

cancer, social well-being and nausea. A total score is obtained. Some limited evidence of discriminative validity is reported from a single evaluation with people with colorectal cancer (Elliot et al., 2004).

The Fox Simple Quality of Life Questionnaire was developed specifically with people with cancer and involved a small number of patients with colorectal cancer as well as lung and ovarian. Good construct content and internal validity is reported (Fox, 2004). No substantial evidence is reported specifically for patients with colorectal cancer.

The Cancer Therapy Satisfaction Questionnaire has been developed for use in a wide range of cancer types and stages but specifically for patients receiving both oral and intravenous chemotherapy, focusing on compliance, feelings about side effects and satisfaction with therapy. There are 18 items but not all are relevant to every patient. Some questions are about oral medication and others for intravenous therapy. Colorectal cancer patients have been involved in the development and subsequent psychometric evaluation; including a UK population. Some supportive evidence of a three domain structure is reported of Side effects, Satisfaction with therapy and Expectations of therapy. Internal consistency is reported as high. Substantial ceiling effects are reported for Expectations of therapy and satisfaction. Moderate correlation of scores is reported between CTSQ domains with corresponding domains of other instruments (Treatment Satisfaction Questionnaire for Medication; EORTC). MID is calculated as 0.5 SD of baseline scores and 1 SEM.

SUMMARY OF EVIDENCE

Generic PROMs

Table 3 summarises the evidence of measurement and operational performance of the two generic PROMs (SF-36 and SF-12) identified in this review applying the adapted appraisal criteria outlined in Appendix B.

Based on the quality of evidence, the SF-12 is clearly the most promising generic health measure compared to the evidence presented for the SF-36. The SF-36 has limited evidence of validity and responsiveness. Moreover, as the SF-12 is in principal a briefer version of the SF-36, it has a shorter completion time which is perhaps also reflected in the instruments high acceptability rate. However, whilst the evidence for the psychometric properties of the SF-12 is strong, support for responsiveness is inconsistent. Therefore, if used with colorectal cancer patients, the SF-12 should be used alongside a condition-specific instrument.

Preference based PROMs

EQ-5D is the only preference-based measure identified by this review. The instrument has some evidence of good psychometric performance and patient acceptability. The EQ-5D also has the advantage of the generation of a utility value.

Cancer-specific PROMs

Three cancer-specific PROMs were identified (EORTC QLQ-C30, EORTC QLQ-CR38 and FACT-C). Table 3 shows the appraisals of the evidence for each of the condition specific PROMs identified in this review.

The EORTC QLQ-C30 is the most evaluated instrument with colorectal cancer patients. All of the evaluations for the EORTC C-30 have been reported in last 4 years; half of those conducted in the UK. The instrument has demonstrated good validity and responsiveness and its easy administration, has proven itself to be especially appropriate within clinical settings. Another notable benefit of the EORTC QLQ-C30 is that the instrument can be applied to any cancer population as it covers aspects of functioning and symptoms that may be of importance to most cancer patients. It is also a brief instrument, consisting of 30 items, which only takes 10-15 minutes to complete. However, it can be argued that the instrument does not cover aspects of health that may be specific to colorectal cancer patients.

Colorectal cancer-specific PROMs

There is a limited volume of evidence for the application of the colorectal cancer specific module of the QLQ-C30 (EORTC QLQ-CR38) and as its application has been inconsistent across studies. Although the instrument appears to be promising in terms of validity, responsiveness and acceptability, the results of these studies should be interpreted with caution as the instrument has not been utilised appropriately as recommended by the developers. For instance, if applied correctly, the QLQ-CR38 would be a 68-item questionnaire which would take approximately 20 minutes to complete, which may have an impact on the psychometric properties of the instrument as well as its acceptability amongst patients. Furthermore, the instrument has now

been replaced and the EORTC developers³ advocate the use of the QLQ-CR29 with colorectal cancer patients. However, there is a distinct lack of evaluations of the QLQ-CR29 with English speaking populations and therefore, the instrument has not been reported in this review.

The FACT-C has also been evaluated with colorectal cancer patients and has shown promising reliability, validity and responsiveness and patient acceptability. The instrument is colorectal cancer-specific and covers five important dimensions (Physical, Social, Emotional and Functional wellbeing as well as colorectal cancer specific Symptoms). The impact of symptoms for patients with colorectal cancer is great. The number of evaluations is similar to those of the EORTC QLQ-C30 and the majority have been conducted in the last 5 years.

The FACT-C is briefer than the EORTC QLQ-C30 and completion time has been reported as 5-10 minutes. However, scoring of the FACT-C is complicated as there are three methods of scoring available for the instrument, each one including a aggregation of different combinations of dimensions. Due to the inconsistency in scoring amongst studies included in this review, it is difficult to provide a comprehensive report on the psychometric properties of the instrument and results should be interpreted with caution.

There is a limited volume of evidence for the application of symptom scales in patients with colorectal cancer; further evaluations are needed.

Survivorship

The instruments included in this review have been used with patients with recent diagnosis or undergoing treatment, principally to evaluate the effectiveness of interventions. However, as more patients are living longer following treatment, there is increasing interest in measuring quality of life amongst the long-term survivors of cancer. In this context, our recommendations regarding generic instruments will likely remain appropriate, but the content of condition-specific instruments may lose face validity as the predicament of survivorship is different to undergoing treatment. There is evidence that the experiences of individuals who have survived cancer for some time are not completely captured by instruments focused on experiences surrounding diagnosis and treatment (Gotay and Muraoka, 1998).

Survival for colorectal cancer has improved considerably and five-year survival estimates for colon and rectal cancer have been reported at 53% and 55%, respectively (National cancer survivorship initiative, 2010). As growing number of patients are living with the long-term effects of treatment whilst many will receive intestinal stoma as part of their treatment. The many HRQoL issues for people with an ostomy include problems with travel, intimacy, and satisfaction with appearance (Krouse et al., 2007). Some studies have documented additional problem areas including sexuality (Symms et al., 2008), psychological well-being (Ries et al., 2008), and interference with work and sport activities (Krouse et al., 2007).

Pearce et al. (2008) identified and appraised several instruments which have been specifically developed for long-term survivors of cancer, including colorectal cancer.

³ This information can be found on the EORTC website <http://groups.eortc.be/qol/qolg_projects.htm>

Currently, from a psychometric perspective, survivorship instruments are deemed to be limited. Avis and Foley (2006) found some positive evidence of responsiveness for a cancer survivor outcome measure, the Quality of Life in Adult Cancer Survivors (QALCAS) Scale. Further consideration would be necessary if the health and well-being of the long-term survivors is to be measured routinely.

Recommendations

There are three main approaches to the measurement of patient-outcomes with people with cancer. The generic approach to health status measurement allows for the comparison across health conditions; the cancer-generic modules which are more focused on dimensions relevant to people with cancer; and add-on modules which focus more on specific domains relative to the type of cancer. This is illustrated in the evidence found in this review.

For the measurement of comprehensive general health status, the SF-12 and the EQ-5D have more psychometric supportive evidence, are acceptable to patients, and should be considered. The EQ-5D is particularly appropriate when a utility score is required.

Based on the volume and quality of evidence, the EORTC QLQ-C30 and the FACT-C are shortlisted as promising cancer-specific PROMs. The more generic QLQ-C30 may be sufficient where the focus is on outcomes for individuals with a broad range of types of cancer, including colorectal cancer.

Table 3: Appraisal of PROMs included in the review

PROM	Reproducibility	Internal consistency	Validity - content	Validity - construct	Responsiveness	Interpretability	Precision	Acceptability	Feasibility
Generic measures									
SF-36	0	0	0	+	+	0	0	0	0
SF-12	0	0	0	++	+	0	0	++	0
Preference-based measures									
EQ-5D	0	n/a	0	++	+	0	0	+	0
Cancer-specific measures									
EORTC C-30	0	0	0	++	++	0	0	++	0
Colorectal cancer-specific measures									
EORTC CR-38	0	0	0	+	+	0	0	+	0
FACT-C	0	++	0	++	++	0	0	+	0

APPENDIX A:

SOURCES FOR SEARCH

Four databases were searched using the search engine Ovid (from January 2006 until February 2010):

- AMED (Allied and Complementary Medicine)
- EMBASE
- Psycinfo
- Ovid MEDLINE (R) In-Process & Other Non-Indexed Citations and Ovid MEDLINE (R)

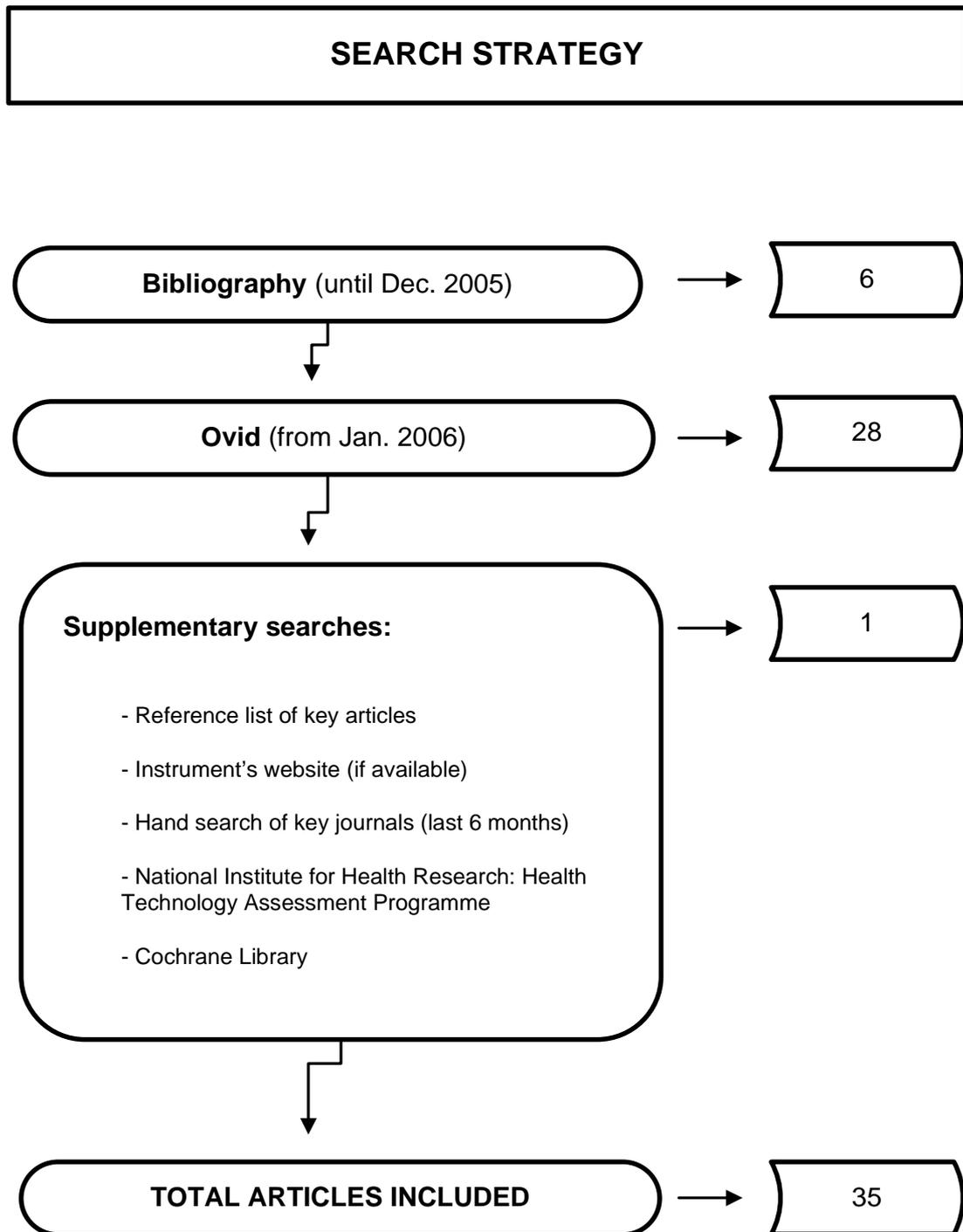
The following is the complete search strategy used to search in databases using the search engine Ovid:

(HR-PRO or HRPRO or HRQL or HRQoL or QL or QoL).ti,ab. or quality of life.mp. or (health index* or health indices or health profile*).ti,ab. or health status.mp. or ((patient or self or child or parent or carer or proxy) adj (appraisal* or appraised or report or reported or reporting or rated or rating* or based or assessed or assessment*)).ti,ab. or ((disability or function or functional or functions or subjective or utility or utilities or wellbeing or well being) adj2 (index or indices or instrument or instruments or measure or measures or questionnaire* or profile or profiles or scale or scales or score or scores or status or survey or surveys)).ti,ab.

and

(("cancer*" or "carcinoma*" or "adenocarcinoma*" or "malignan*" or "tumor*" or "tumour*" or "neoplasm*" or "metasta*") adj5 (colorectal or colon or colonic or rectal or rectum or rectosigmoid or bowel)).ab,ti.

SEARCH STRATEGY



APPENDIX B: Appraisal of the methodological quality of PROMs

A simple rating scale (Table i) was used to rate the sum total of evidence available for each dimension or criterion against which PROMs were assessed. The dimensions or criteria are summarised in Table ii.

Table i: Psychometric and operational criteria

0	<i>not reported (no evaluation completed)</i>
—	<i>Evaluation evidence available indicating poor performance of instrument</i>
+	<i>Some limited evidence in favour</i>
++	<i>Good evidence in favour</i>
+++	<i>Excellent evidence in favour</i>

Appraisal criteria (adapted from Smith et al., 2005 and Fitzpatrick et al., 1998; 2006)

Appraisal component	Definition/test	Criteria for acceptability
Reliability		
Test-retest reliability	The stability of a measuring instrument over time; assessed by administering the instrument to respondents on two different occasions and examining the correlation between test and re-test scores	Test re-test reliability correlations for summary scores 0.70 for group comparisons
Internal consistency	The extent to which items comprising a scale measure the same construct (e.g. homogeneity of items in a scale); assessed by Cronbach's alpha's and item-total correlations	Cronbach's alphas for summary scores ≥ 0.70 for group comparisons Item-total correlations ≥ 0.20
Validity		
Content validity	The extent to which the content of a scale is representative of the conceptual domain it is intended to cover; assessed qualitatively during the questionnaire development phase through pre-testing with patients. Expert opinion and literature review	Qualitative evidence from pre-testing with patients, expert opinion and literature review that items in the scale represent the construct being measured Patients involved in the development stage and item generation
Construct validity	Evidence that the scale is correlated with other measures of the same or similar constructs in the hypothesised direction; assessed on the basis of correlations between the measure and other similar measures	High correlations between the scale and relevant constructs preferably based on a priori hypothesis with predicted strength of correlation
	The ability of the scale to differentiate known-groups; assessed by comparing scores for sub-groups who are expected to differ on the construct being measured (e.g a clinical group and control group)	Statistically significant differences between known groups and/or a difference of expected magnitude
Responsiveness	The ability of a scale to detect significant change over time; assessed by comparing scores before and after an intervention of known efficacy (on the basis of various methods including t-tests, effect sizes (ES), standardised response means (SRM) or responsiveness statistics	Statistically significant changes on scores from pre to post-treatment and/or difference of expected magnitude
Floor/ceiling effects	The ability of an instrument to measure accurately across full spectrum of a construct	Floor/ceiling effects for summary scores $<15\%$
Practical properties		
Acceptability	Acceptability of an instrument reflects respondents' willingness to complete it and impacts on quality of data	Low levels of incomplete data or non-response
Feasibility/burden	The time, energy, financial resources, personnel or other resources required of respondents or those administering the instrument	Reasonable time and resources to collect, process and analyse the data

APPENDIX C: Generic and Utility PROMs

a) SF-36: Medical Outcomes Study 36-item Short Form Health Survey (Ware and Sherbourne, 1992; Ware et al., 1994; Ware, 1997)

The Medical Outcomes Study (MOS) Short Form 36-item Health Survey (SF-36) is derived from the work of the Rand Corporation during the 1970s (Ware and Sherbourne, 1992; Ware et al., 1994; Ware, 1997). It was published in 1990 after criticism that the SF-20 was too brief and insensitive. The SF-36 is intended for application in a wide range of conditions and with the general population. Ware et al., (1994; 1997) proposed that the instrument should capture both mental and physical aspects of health. International interest in this instrument is increasing, and it is by far the most widely evaluated measure of health status (Garratt et al., 2002).

Items were derived from several sources, including extensive literature reviews and existing instruments (Ware and Sherbourne, 1992; Ware and Gandek, 1998; Jenkinson and McGee, 1998). The original Rand MOS Questionnaire (245 items) was the primary source, and several items were retained from the SF-20. The 36 items assess health across eight domains (Ware, 1997), namely Bodily pain (BP: two items), General health perceptions (GH: five items), Mental health (MH: five items), Physical functioning (PF: ten items), Role limitations due to emotional health problems (RE : three items), Role limitations due to physical health problems (RP: four items), Social functioning (SF: two items), and Vitality (VT: four items). An additional health transition item, not included in the final score, assesses change in health

There are between two and six categorical response options for each item. Scoring uses a weighted scoring algorithm and a computer-based programme is recommended. Eight domain scores give a health profile; scores are transformed into a scale from 0 to 100, where 100 denotes the best health. Scores can be calculated when up to half of the items are omitted. Two component summary scores for physical and mental health (MPS and MCS, respectively) can also be calculated. A version of the SF-36 plus three depression questions has been developed and is variously called the Health Status Questionnaire (HSQ) or SF-36-D.

The SF-36 can be self-, interview-, or telephone-administered.

b) SF-12: Medical Outcomes Study 12-item Short Form Health Survey (Ware et al., 1995)

In response to the need to produce a shorter instrument that could be completed more rapidly, the developers of the Medical Outcomes Study (MOS) 36-item Short Form Health Survey (SF-36) produced the 12-item Short Form Health Survey (SF-12) (Ware et al., 1995). Using regression analysis, 12 items were selected that reproduced 90% of the variance in the overall Physical and Mental Health components of the SF-36 (Table 3.1). The same eight domains as the SF-36 are assessed and categorical response scales are used. A computer-based scoring algorithm is used to calculate scores: Physical Component Summary (PCS) and Mental (MCS) Component Summary scales are generated using norm-based methods. Scores are transformed to have a mean value of 50, standard deviation (SD) 10, where scores above or below 50 are above or below average physical or mental well-being, respectively. Although not recommended by the developers, Schofield and Mishra (1998) report eight domain scores and two summary scores. The SF-12 may be self-, interview-, or telephone-administered.

Several authors have proposed simplification of the scoring process and revision of the SF-12 summary score structure, where norm-based weighting is replaced by item summation to facilitate score interpretation (Resnick and Nahm, 2001; Resnick and Parker, 2001).

c) European Quality of Life instrument, EQ-5D (The EuroQol Group, 1990; revised 1993)

The European Quality of Life instrument (EuroQol), now generally known as the EQ-5D, was developed by researchers in five European countries to provide an instrument with a core set of generic health status items (The EuroQol Group, 1990; Brazier et al., 1993). Although providing a limited and standardized reflection of HRQoL, it was intended that use of the EuroQol would be supplemented by disease-specific instruments. The developers recommend the EuroQol for use in evaluative studies and policy research; given that health states incorporate preferences, it can also be used for economic evaluation. It can be self or interview-administered.

Existing instruments, including the Nottingham Health Profile, Quality of Well-Being Scale, Rosser Index, and Sickness Impact Profile were reviewed to inform item content (The EuroQol Group, 1990). There are two sections to the EuroQol: the EQ-5D and the EQ thermometer. The EQ-5D assesses health across five domains: anxiety/depression (AD), mobility (M), pain/discomfort (PD), self-care (SC), and usual activities (UA). Each domain has one item and a three-point categorical response scale; health 'today' is assessed. Weights based upon societal valuations of health states are used to calculate an index score of -0.59 to 1.00 , where -0.59 is a state worse than death and 1.00 is maximum well-being. A score profile can be reported. The EQ thermometer is a single 20cm vertical visual analogue scale with a range of 0 to 100, where 0 is the worst and 100 the best imaginable health.

GENERIC INSTRUMENTS DOMAINS AND SCORING

Table 5: Domains and Scoring

<i>Instrument</i>	<i>Domains (No. Items)</i>	<i>Response Options</i>	<i>Score</i>	<i>Completion</i>
SF-36	Bodily pain (BP) (1), General health (GH) (5) Mental health (MH) (5), Physical functioning (PF) (6) Role functioning (RF), Social functioning (SF) (1)	Categorical: 2-6 options Recall: standard 4 weeks, acute 1 week	Algorithm Domain profile (0-100, 100 best health) Summary: Physical (PCS) and Mental (MCS) (mean 50, sd 10).	Interview or self. 5-10 minutes.
SF-12	Bodily pain (BP) (1), Energy/Vitality (VT) (1), General health (GH) (1), Mental health (MH) (2), Physical functioning (PF) (2), Role limitation-emotional (RE) (2), Role limitation-physical (RP) (2), Social functioning (SF) (1)	Categorical: 2-6 options Recall: standard 4 weeks, acute 1 week	Algorithm Domain profile (0-100, 100 best health) Summary: Physical (PCS), Mental (MCS) (mean 50, sd 10)	Interview or self
EQ-5D	Anxiety/depression (1), mobility (1), pain/discomfort (1), self-care (1), usual activities (1). EQ-thermometer: Global health (1)	Categorical: 3 options - 1 (no problems), 2 (some problems), and 3 (inability/extreme problems), EQ-thermometer: VAS current health.	Summation: responses are scored for each individual health state and then multiplied by the duration of time in that health state; Utility index (-0.59-100, where -0.59 is a state worse than death and 1.00 is maximum well-being. EQ-thermometer: VAS (0-100, 100 representing best imaginable health).	Interview or self.

Table 6: Summary of generic instruments: health status domains (after Fitzpatrick et al., 1998)

<i>Instrument</i>	<i>Instrument domains</i>							<i>Personal constructs</i>
	<i>Physical function</i>	<i>Symptoms</i>	<i>Global judgement of health</i>	<i>Psychological well-being</i>	<i>Social well-being</i>	<i>Cognitive functioning</i>	<i>Role activities</i>	
SF-36	x	x	x	x	x		x	
SF-12	x	x	x	x	x		x	
EQ-5D	x	x	x	x	x		x	

APPENDIX D: Cancer-specific PROMs

European Organization for Research and Treatment of Cancer Quality of Life

Questionnaire (EORTC QLQ-C30)

The EORTC QLQ-C30 (Aaronson et al., 1993) is a 30-item questionnaire composed of five multi-item functional subscales: physical, role, emotional, social, and cognitive functioning; three multi-item symptom scales: fatigue, pain, and emesis; global health/QoL subscale; and six single items to assess financial impact and symptoms such as dyspnoea, sleep disturbance, appetite, diarrhoea, and constipation. The questionnaire employs 28 four-point Likert scales of 'not at all' to 'very much,' and two seven-point Likert scales for the global health and QoL domain, total scores ranging from 0 to 100. For functional and global QoL scales, higher scores represent a better level of functioning. For symptom-oriented scales, a higher score represents more severe symptoms. Extensive patient input during an international field study contributed to the development of the QLQ-C30, initially with lung cancer patients (Aaronson et al., 1993) and subsequently with patients with heterogeneous diagnoses (Osoba et al., 1994). Content validity of the QLQ-C30 has been maintained via modifications to improve the content, specifically in terms of the Role Functioning scale and a conceptual difficulty (undue emphasis on physical functioning) in the global QoL scale (Osoba et al., 1997).

European Organization for Research and Treatment of Cancer Colorectal Cancer Specific Quality of Life Questionnaire (EORTC QLQ-CR38)

The EORTC QLQ-C38 (Sprangers et al., 1999) is a 38-item colorectal cancer-specific module to be administered in addition to the EORTC QLQ-C30 questionnaire. The 38 items are added to the 30 core items of the QLQ-C30, becoming a 68 item questionnaire. The 38 item module consists of both multi-item scales and single-item measures. Including the domains of the QLQ-C30, the QLQ-CR38 adds two functional scales (body image (BI) and sexual function (SF)), seven symptom scales (micturition problems (MP), gastrointestinal tract symptoms (GT), chemotherapy side effects (GS), defecation problems (DP), stoma-related problems (SP), and male and female sexual problems (MS & FS)), and three single-item measures (sexual enjoyment, weight loss, and future perspective). Items are scored on a four-point Likert scales of 1 (not at all) to 4 (very much). The values in the functional scales and items range from 0 (worse outcome) to 100 (better outcome), whereas those in the symptom scales or items range from 0 (fewer symptoms) to 100 (more symptoms).

Functional Assessment of Cancer Therapy –Colorectal (FACT-C)

The FACT-C is a 7 item colorectal cancer specific module that supplements the Functional Assessment of Cancer Therapy – General (FACT-G). The 10 items are added to the 27 core items of the FACT-G, becoming a 37 item questionnaire. The FACT-C consists of 5 subscales: physical wellbeing (PW, 7 items), social and family wellbeing (SW, 7 items), emotional wellbeing (EW, 6 items), functional wellbeing (FW, 7 items) and the Colorectal Cancer Subscale (CCS, 10 items). Scores can be produced through three different calculations; a combined total of all domains (FACT-C total), the Colorectal Cancer Score (CCS) and a Treatment Outcome Index (TOI) can be calculated by summing the FACT-G physical and functional domains and the CCS. Answers are provided on a scale of Not at all; A little; Somewhat; Quite a bit; Very Much, in the context of 'during the past seven days'. Items are scored from 0-4, with negatively-phrased items requiring reverse response scores. Higher scores represent better well-being on each of the dimensions or better global QoL when combined.

CANCER-SPECIFIC INSTRUMENTS DOMAINS AND SCORING

Table 7: Domains and Scoring

<i>Instrument name (total no. items)</i>	<i>Domains (no. items)</i>	<i>Response options</i>	<i>Scoring</i>	<i>Mode of administration Completion time</i>	<i>Licensing information</i>
European Organization for Research and Treatment of Cancer Quality of Life core Questionnaire, EORTC QLQ-C30 (30)	Global health status/QoL (2) <i>Functional scales:</i> Physical functioning (5) Role functioning (2) Emotional functioning (4) Cognitive functioning (2) Social functioning (2) <i>Symptom scales:</i> Fatigue (3) Nausea & vomiting (2) Pain (2) Dyspnoea (1) Insomnia (1) Appetite loss (1) Constipation (1) Diarrhoea (1) Financial difficulties (1)	4-point Likert scales where 1=not at all (best), 4=very much (worst). 7-point Likert scales for global health and overall QoL questions, where 1=very poor, 7=excellent Except for PF, responses are based on the past week	Subscale scores transformed into 0-100 scores using an algorithm. Higher scores on functional scales and global items indicate better functioning; higher scales on symptom scales indicate worse symptomatology. Aggregation of subscale scores not recommended by developers.	Interview, telephone, or self-administration. Electronic versions of the QLQ-C30 under development. 10-15 minutes	No charge for use in academic settings, but written consent required for each study. Royalty fee, based on no. of patients, payable for commercial studies.

<p>European Organization for Research and Treatment of Cancer - Colorectal cancer module, EORTC QLQ-CR38 (30+38) <i>to be used in conjunction with EORTC QLQ-C30</i></p>	<p>EORTC QLQ-C30 items (30) <i>As above</i></p> <p>additional items (38) Micurition problems GI tract symptoms Chemotherapy side-effects Defecation problems Stoma-related problems Sexual problems/enjoyment Weight loss Future perspective</p>	<p>As above</p>	<p>As above</p>	<p>Interview, telephone, or self-administration. Electronic versions of the QLQ-C30 under development.</p> <p>15-20 minutes</p>	<p>As above</p>
<p>Functional Assessment of Cancer Therapy – Colorectal cancer, FACT-C (37)</p>	<p>FACT-G items (27) Physical well-being, symptoms (7) Social/family well-being (7) Emotional well-being (6) Functional well-being, personal constructs (7)</p> <p>Additional concerns (10) GI symptoms (4) Appetite (1) Weight loss (1) Body image (1) Ostomy appliance (yes/no)? Oa – stigma (1) Oa – bother (1)</p>	<p>5-point Likert scales, where 0=not at all, 4=very much</p> <p>Responses based on the past seven days</p>	<p>Multiple scoring options: subscale, overall total, Trial Outcome Index.</p>	<p>Interview, telephone, or self-administration. Computerised modes of administration under development</p> <p>5-10 minutes</p>	<p>Use of English versions of FACT/FACIT measures is free of charge, on condition of sharing data. Users must complete an agreement and submit project information for each study.</p>

Table 8: Summary of cancer-specific instruments: health status domains (after Fitzpatrick et al., 1998)

<i>Instrument</i>	<i>Instrument domains</i>								
	Physical function	Symptoms	Global judgement of health	Psychological well-being	Social well-being	Cognitive functioning	Role activities	Personal constructs	Satisfaction with care
EORTC QLQ-C30	x	x	x	x	x	x	x	x	
EORTC QLQ-C30 & EORTC QLQ-CR38*	x	x	x	x	x	x	x	x	
FACT-C	x	x	x	x	x		x	x	

* crosses in bold indicate content of the colorectal cancer-specific module; the remainder indicate domains covered by the core

APPENDIX E: LICENSING & CONTACT DETAILS

European Organization for Research and Treatment of Cancer Quality of Life Questionnaire (EORTC QLQ-C30)

EORTC Headquarters

Quality of Life Department

Ave. E. Mounier 83, B.11

1200 Brussels Belgium

Fax: +32 (0)2 779 45 68

Tel: +32 (0)2 774 1678

E mail: ken.cornelissen@eortc.be

http://groups.eortc.be/qol/questionnaires_qlqc30.htm

http://groups.eortc.be/qol/questionnaires_modules.htm

Functional Assessment of Cancer Therapy – General Version (FACT-G)

FACIT.org

381 South Cottage Hill Ave

Elmhurst, IL 60126 USA

Tel: (+1) 877 828 3228

Fax: (+1) 630 279 9465

E mail: information@facit.org

<http://www.facit.org/qview/qlist.aspx>

REFERENCES

- 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. *Journal of the National Cancer Institute* 1993; 85(5):365-376.
- Au HJ, Karapetis CS, O'Callaghan CJ *et al.* Health-related quality of life in patients with advanced colorectal cancer treated with cetuximab: overall and KRAS-specific results of the NCIC CTG and AGITG CO.17 Trial. *Journal of Clinical Oncology* 2009; 27:1822-1828.
- Bailey C, Corner J, Addington HJ, Kumar D, Haviland J. Older patients' experiences of treatment for colorectal cancer: an analysis of functional status and service use. *European-Journal-of-Cancer-Care* 2004; 13:483-493.
- Best J, Garrison L, Hollingworth W, Ramsey S, Veenstra D. Preference values associated with stage III colon cancer and adjuvant chemotherapy. *Quality of Life Research* 2010; 19:391-400.
- Brazier JE, Jones N, Kind P. Testing the validity of the EuroQol and comparing it with the SF-36 Health Survey questionnaire. *Quality of Life Research* 1993; 2(3):169-180.
- Bruera E, Kuehn N, Miller MJ, Selmsler P, Macmillan K. The Edmonton Symptom Assessment System (ESAS): a simple method of the assessment of palliative care patients. *Journal of Palliative Care* 1991; 7:69
- Cella DF, Tulsky DS, Gray G *et al.* The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *Journal of Clinical Oncology* 1993; 11:570-579.
- Chang VT, Hwang SS, Kasimis B, Thaler HT. Shorter symptom assessment instruments: the Condensed Memorial Symptom Assessment Scale (CMSAS). *Cancer Invest* 2004; 22:526-536.
- Cleeland CS, Mendoza TR, Wang XS, Chou C, Harle M, Morrissey M, & Engstrom MC. Assessing Symptom Distress in Cancer: The M. D. Anderson Symptom Inventory. *Cancer* 2000; 89:1634-1646
- Darzi A. Our NHS Our Future: NHS Next Stage Review Interim Report. Department of Health, London, October 2007.
- Darzi A. Our NHS Our Future: High Quality Care for All. NHS Next Stage Review Final Report. Department of Health, London, June 2008.
- Department of Health, London, December 2007. NHS Cancer Reform Strategy.
- Department of Health, London, December 2008. Guidance on the routine collection of Patient Reported Outcome Measures (PROMs) for the NHS in England 2009/10.

Elliott BA, Renier CM, Haller IV, Elliott TE. Health-related quality of life (HRQoL) in patients with cancer and other concurrent illnesses. *Quality of Life Research* 2004; 13:457-462.

EuroQol Group, The. EuroQol: a new facility for the measurement of health-related quality of life. *Health Policy* 1990; 16(3):199-208.

Fitzpatrick R, Davey C, Buxton MJ, Jones DR. Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment* 1998; 2(14).

Fox SW. Preliminary psychometric testing of the Fox Simple Quality of Life Scale. *Journal of Neuroscience Nursing* 2004; 36:157-166.

Gall CA, Weller D, Esterman A et al. Patient satisfaction and health-related quality of life after treatment for colon cancer. *Diseases of the Colon & Rectum* 2007; 50:801-809.

Garratt AM, Schmidt L, Mackintosh A, Fitzpatrick R. Quality of life measurement: bibliographic study of patient assessed health outcome measures. *British Medical Journal* 2002; 324 (7351):1417-1421.

Gotay C, Muraoka M, Quality of life in long term survivors of adult onset cancers. *Journal of the National Cancer Institute*. 1998; 90:656-657

Gupta D, Lis CG, Granick J, Grutsch JF, Vashi PG, Lammersfeld CA. Malnutrition was associated with poor quality of life in colorectal cancer: a retrospective analysis. *Journal of Clinical Epidemiology* 2006; 59:704-709.

Hassan I, Larson DW, Cima RR et al. Long-term functional and quality of life outcomes after coloanal anastomosis for distal rectal cancer. *Diseases of the Colon & Rectum* 2006; 49:1266-1274.

Jayne DG, Guillou PJ, Thorpe H et al. Randomized trial of laparoscopic-assisted resection of colorectal carcinoma: 3-year results of the UK MRC CLASICC Trial Group. *Journal of Clinical Oncology* 2007; 25:3061-3068.

Jenkinson C, McGee H. Health Status Measurement, a Brief but Critical Introduction. Oxford: Radcliffe Medical Press Ltd, 1998.

Jonker DJ, O'Callaghan CJ, Karapetis CS et al. Cetuximab for the treatment of colorectal cancer. *New England Journal of Medicine* 2007; 357:2040-2048.

Kabbinavar FF, Wallace JF, Holmgren E et al. Health-related quality of life impact of bevacizumab when combined with irinotecan, 5-fluorouracil, and leucovorin or 5-fluorouracil and leucovorin for metastatic colorectal cancer. *Oncologist* 2008; 13:1021-1029.

Kariv Y, Kariv R, Hammel JP, Lavery IC. Postoperative radiotherapy for stage IIIA rectal cancer: is it justified? *Diseases of the Colon & Rectum* 2008; 51:1459-1466.

King PM, Blazeby JM, Ewings P et al. The influence of an enhanced recovery programme on clinical outcomes, costs and quality of life after surgery for colorectal cancer. *Colorectal Disease* 2006; 8:506-513.

King PM, Blazeby JM, Ewings P, Kennedy RH. Detailed evaluation of functional recovery following laparoscopic or open surgery for colorectal cancer within an enhanced recovery programme. *International Journal of Colorectal Disease* 2008; 23:795-800.

Kirkova Jordanka; Davis Mellar P; Walsh Declan; Tiernan Eoin; O'Leary Norma; LeGrand Susan B; Lagman Ruth L; Russell K Mitchell Cancer symptom assessment instruments: a systematic review. *Journal of clinical oncology*. 2006; 24(9):1459-73.

Knowles G, Sherwood L, Dunlop MG et al. Developing and piloting a nurse-led model of follow-up in the multidisciplinary management of colorectal cancer. *European Journal of Oncology Nursing* 2007; 11:212-223.

Krouse RS, Grant M, Wendel CS, et al: A mixed-methods evaluation of health-related quality of life for male veterans with and without intestinal ostomy. *Disease of the Colon and Rectum* 2007; 50:2054-2066.

Lis CG, Gupta D, Granick J, Grutsch JF. Can patient satisfaction with quality of life predict survival in advanced colorectal cancer? *Supportive Care in Cancer* 2006; 14:1104-1110.

Lis CG, Gupta D, Grutsch JF. Using the quality of life index as a prognostic stratification tool in advanced cancer. *Journal of Clinical Outcomes Management* 2007; 14:144-148.

Mastracci TM, Hendren S, O'Connor B, McLeod RS. The impact of surgery for colorectal cancer on quality of life and functional status in the elderly. *Diseases of the Colon & Rectum* 2006; 49:1878-1884.

McMillan cancer Support. *National cancer survivorship initiative: Vision*. [London]: Department of Health 2010.

Neuman HB, Schrag D, Cabral C et al. Can differences in bowel function after surgery for rectal cancer be identified by the European Organization for Research and Treatment of Cancer quality of life instrument? *Annals of Surgical Oncology* 2007; 14:1727-1734.

Osoba D, Zee B, Pater JL, Warr D, Kaizer L, Latreille J. Psychometric properties and responsiveness of the EORTC Quality of Life Questionnaire (QLQ-C30) in patients with breast, ovarian, and lung cancer. *Quality of Life Research* 1994; 3(5):353-64.

Osoba D, Aaronson NK, Zee B, Sprangers MAG, Te Velde A. Modification of the EORTC QLQ-C30 (version 2.0) based on content validity and reliability testing in large samples of patients with cancer. *Quality of Life Research* 1997; 6(2):103-108.

- Osoba D, Rodrigues G, Myles J, Zee B, Pater JL. Interpreting the significance of changes in health-related quality-of-life scores. *Journal of Clinical Oncology* 1998; 16(1):139-144.
- Pearce N.J, Sanson-Fisher R, & Campbell HS. Measuring quality of life in cancer survivors: a methodological review of existing scales. *Psycho-oncology* 2008; 17(7):629-640.
- Portenoy RK, Thaler HT, Kornblith AB et al. The Memorial Symptom Assessment Scale: an instrument for the evaluation of symptom prevalence, characteristics and distress. *European Journal of Cancer, Part A: General Topics* 1994; 30A:1326-1336.
- Reeve BB, Potosky AL, Smith AW et al. Impact of cancer on health-related quality of life of older Americans. *Journal of the National Cancer Institute* 2009;101:860-868.
- Sharma A, Sharp DM, Walker LG, Monson JR. Predictors of early postoperative quality of life after elective resection for colorectal cancer. *Annals of Surgical Oncology* 2007; 14:3435-3442.
- Resnick B, Nahm ES. Reliability and validity testing of the revised 12-item Short-Form Health Survey in older adults. *Journal of Nursing Measurement* 2001; 9(2):151-61.
- Resnick B, Parker R. Simplified scoring and psychometrics of the revised 12-item Short-Form Health Survey. *Outcomes Management for Nursing Practice* 2001; 5(4): 161-6.
- Ries LAG, Merbert D, Krapcho M, et al (eds): Surveillance, Epidemiology, and End results (SEER) Cancer Statistics Review, 1975-2005. *Bethesda, MD, National Cancer Institute* 2008. http://seer.cancer.gov/csr/1975_2005/
- Schofield M, Mishra G. Validity of the SF-12 compared with the SF-36 health survey in pilot studies of the Australian longitudinal study on women's health. *Journal of Health Psychology* 1998; 3(2):259-271.
- Simon AE, Thompson MR, Flashman K, Wardle J. Disease stage and psychosocial outcomes in colorectal cancer. *Colorectal Disease* 2009; 11:19-25.
- Sprangers MAG, Te Velde A, Aaronson NK. The construction and testing of the EORTC colorectal cancer-specific quality of life questionnaire module (QLQ-CR38). *European Journal of Cancer* 1999; 35 (2):238-247.
- Steginga SK, Lynch BM, Hawkes A, Dunn J, Aitken J. Antecedents of domain-specific quality of life after colorectal cancer. *Psycho-Oncology* 2009; 18:216-220.
- Stephenson LE, Bebb DG, Reimer RA, Culos-Reed SN. Physical activity and diet behaviour in colorectal cancer patients receiving chemotherapy: associations with quality of life. *BMC Gastroenterology* 2009; 9:60.
- Symms MR, Rawl SM, Grant M, et al: Sexual health and quality of life among male veterans with intestinal ostomies. *Clinical Nurse Specialist* 2008; 22:30-40.

Trask PC, Tellefsen C, Espindle D, Getter C, Hsu MA. Psychometric validation of the cancer therapy satisfaction questionnaire. *Value in Health* 2008; 11:669-679.

Twelves C, Gollins S, Grieve R, Samuel L. A randomised cross-over trial comparing patient preference for oral capecitabine and 5-fluorouracil/leucovorin regimens in patients with advanced colorectal cancer. *Annals of Oncology* 2006; 17:239-245.

Ward WL, Hahn EA, Mo F, Hernandez L, Tulskey DS, Cella DF. Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrumentality. *Quality of Life Research* 1999; 8:181-195.

Ware JE, Sherbourne CD. The MOS 36-Item Short-Form Health Survey (SF-36). I. Conceptual framework and item selection. *Medical Care* 1992; 30(6):473-483.

Ware J, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary Scales: A User's Manual. Boston, MA: The Health Institute, New England Medical Centre, 1994.

Ware JE, Kosinski M, Keller SD. SF-12: How to Score the SF-12 Physical and Mental Health Summary Scales. Boston, MA: The Health Institute, New England Medical Centre, 1995. 2nd edition.

Ware JE. SF-36 Health Survey. Manual and Interpretation Guide. Boston, MA: The Health Institute, New England Medical Centre, 1997. 2nd edition.

Ware JE, Gandek B. Methods for testing data quality, scaling assumptions and reliability: the IQOLA project approach. *Journal of Clinical Epidemiology* 1998; 51(11):945-952.

Ware JE, Kosinski M, Dewey JE *et al.* How to score and interpret single-item health status measures: a manual for users of the SF-8 Health Survey. Lincoln, RI, USA: Quality Metric Incorporated, 2001.

Whistance RN, Conroy T, Chie W *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. *European Journal of Cancer* 2009; 45:3017-3026.

Wilson TR, Alexander DJ, Kind P. Measurement of health-related quality of life in the early follow-up of colon and rectal cancer. *Diseases of the Colon & Rectum* 2006; 49:1692-1702.

Wilson TR, Alexander DJ. Clinical and non-clinical factors influencing postoperative health-related quality of life in patients with colorectal cancer. *British Journal of Surgery* 2008; 95:1408-1415.

Yau T, Watkins D, Cunningham D, Barbachano Y, Chau I, Chong G. Longitudinal assessment of quality of life in rectal cancer patients with or without stomas following primary resection. *Diseases of the Colon & Rectum* 2009; 52:669-677.

Yost KJ, Cella D, Chawla A et al. Minimally important differences were estimated for the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) instrument using a combination of distribution- and anchor-based approaches. *J Clin Epidemiol* 2005; 58:1241-1251.