Running title: Quality of CRC HRQOL comparative studies

Title: Methodological and reporting quality of comparative studies evaluating health-related quality of life of colorectal cancer patients and controls: A systematic review

Correspondence Author:

Name: Carlos King Ho Wong, PhD, MPhil, BSc

Institution: Department of Family Medicine and Primary Care, The University of Hong Kong

Address: 3/F, Ap Lei Chau Clinic, 161 Ap Lei Chau Main Street, Ap Lei Chau, Hong Kong

Contact: +852-25185688 (tel); +852-28147475 (fax) carlosho@hku.hk (email)

Order of Author: Carlos K.H. Wong PhD*[1] , Vivian Y. W. Guo[1], Jing Chen PhD[2], Cindy L. K. Lam[1]

[*] First and correspondence Author

[1] Department of Family Medicine and Primary Care, The University of Hong Kong

[2] School of Nursing, The University of Hong Kong

Conflicts of Interest and Source of Funding: None were declared.

Author contributions:

CKHW: conception and design, acquisition of data, analysis and interpretation of data; drafting and revising the article. VYWG: analysis and interpretation of data, revising it critically for important intellectual content. JC: acquisition of data, analysis and interpretation of data. CLKL: interpretation of data and revising the article. All authors approved the final version of manuscript.

Word count for the text: 3646

Word count for the abstract: 300

Abstract

Background: Health-related quality-of-life is an important outcome measure in patients with colorectal cancer. Comparison with normative data is increasingly undertaken to assess the additional impact of colorectal cancer on health-related quality of life.

Objective: This review aimed to critically appraise methodological details and reporting characteristics of comparative studies evaluating differences in health-related quality-of-life between patients and controls.

Data sources: A systematic search of English-language literature published between January 1985 and May 2014 was conducted through a database search of Pubmed, Web of Science, Embase, and Medline.

Study Selection: Comparative studies reporting health-related quality-of-life outcomes among colorectal cancer patients and controls.

Main Outcome Measures: Methodological and reporting quality per comparison study was evaluated on a 11-item methodological checklist proposed by Efficace and a set of criteria pre-determined by reviewers.

Results: Thirty-one comparative studies involving >10,000 patients and >10,000 controls were included. Twenty-three studies (74.2%) originated from European countries, with the largest number from Netherland (n=6). Twenty-eight studies (90.3%) compared health-related quality-of-life of patients with normative data published elsewhere, while the remaining recruited a group of colorectal cancer patients and a group of control patients within the same studies. The EORTC QLQ-C30 was the most extensively used instrument (n=16; 51.6%). Eight studies (25.8%) were classified as "probably robust" for clinical decision making according to the Efficace's standard methodological checklist. Our further quality assessment revealed the lack of scores differences reported (61.3%), contemporary comparisons (36.7%), statistical significance tested (38.7%) and matching of control group (58.1%), possibly leading to inappropriate control groups for fair comparisons.

Limitations: Meta-analysis of differences between the two groups was not available.

Conclusions: One-fourth of comparative studies evaluating health-related quality-of-life of colorectal cancer generally achieved high-quality in reporting characteristics and methodological details. Future studies are encouraged to undertake health-related quality-of-life measurement and adhere with methodological checklist when compared to controls.

Keywords: systematic review; comparative study; colorectal cancer; quality of life; normative

Manuscript Text

Introduction

Colorectal cancer (CRC) is one of the commonest cancers and leading cause of cancer deaths worldwide[1]. Due to the emerging treatment therapies for CRC, significant group of CRC patients survived with prolonged life expectancy[2], whose health-related quality of life (HRQOL) outcomes were of interest to the clinicians and decision-makers for the development of optimal treatment strategies inducing preservation of HRQOL[3]. Furthermore, the importance of incorporating HRQOL assessments in oncological clinical trials has been well recognized in comparative effectiveness research to aid in clinical practice and decision making[4].

Comparisons with published normative data are undertaken to assess the additional impact of cancer and cancer treatment on HRQOL scores[5, 6], given the co-existence of chronic conditions affecting HRQOL likely. Such comparison is important because norm-based comparison allows for quantifying the extent of departures from the norm, and facilitating interpretation of the clinical importance of HRQOL scores. With reference to country-specific normative data, HRQOL of CRC patients were increasingly compared with that of non-CRC control group as an indication of the HRQOL restriction to CRC patients. Most studies reported HRQOL outcomes comparisons between the CRC and non-CRC control groups with reference to country-specific normative data. Findings from a systematic review of ten HRQOL studies among long-term CRC survivors[7] concluded that CRC patients appeared to have comparable psychological aspect of HRQOL but slightly lower physical aspect of HRQOL than available normative data. Of note, the methodological standards and

normative data for such HRQOL comparisons varied across studies[7, 8], hampering the value and importance of clinical interpretation.

In spite of numerous studies available over the past two decades, there has been an increased concern on the methodological and reporting quality of HRQOL studies in clinical trials involving CRC patients. A 11-item checklist proposed by Efficace et al.[9] has been widely applicable to evaluate the quality of HRQOL reporting in oncological clinical trials. Particularly for CRC studies, systematic review[10] identified methodological shortcomings of thirty-one randomized controlled trials measuring HRQOL as primary or secondary endpoints, particularly addressing the lack of baseline compliance and missing data reported in a majority of studies. However, evidence on detailed methodological critique of quality and reporting characteristics of comparative studies that assessed HRQOL differences between groups is limited. The purpose of this study was to conduct a systematic review of the methodological quality of comparative studies that compared HRQOL between CRC and control groups by either generic or cancer-specific instruments, including assessment of the extent to which the specific domains of HRQOL of CRC patients was significant different from that of healthy controls.

Methods

Search Engine and Strategy

Systematic literature search was conducted on May 2014 in databases of PubMed, Web of Science using Web of Knowledge platform, Embase and MEDLINE using OVID

searching platform, to identify studies that investigated the HRQOL of colorectal neoplasm patients. The Medical Subject Heading (MESH) 'quality of life' was combined with 'colon neoplasm', 'colon cancer', 'rectal cancer', 'rectal neoplasm' and 'colorectal cancer'. Studies were limited to English language, and the years between January 1985 and March 2014. Electronic search strategy in each electronic database is showed in Appendix, and has been adopted in one previous systematic review[11]. No additional hand search was done.

After the initial check for duplicated articles, the titles and abstracts of remaining articles were screened to rule out the introductories, editorials, letters, commentaries, study protocols, case reports, pure literature reviews and meta-analyses, conference proceedings, past and current clinical guidelines and recommendations. Selected articles were further screened with full texts. The eligibility criteria of studies were 1) to involve original articles, 2) to measure HRQOL using standardized instruments with items rating on point Likert scales or on linear analogue scales, and 3) to compare at least one HRQOL outcome between CRC patients and the general population/healthy controls. Articles without available full-text were excluded. Articles were also excluded if no abstract and full text available. When there were multiple reports of studies using the same sample, the most updated publication of a study was included.

Two reviewers (CW and CJ) independently screened the eligibility criteria of study titles, abstracts, and selected full-texts of the studies retrieved by the literature search. Thus, assessment of the quality in eligible studies was performed by two reviewers independently (CW and CJ). Disagreements regarding the procedures of database search, study selection and eligibility were resolved by discussion.

Data Extraction

Study characteristics including first author, year of publication, country of origin, study design, population, sample size, demographics of CRC patients and non-CRC control population, response rate, HRQOL instruments, and HRQOL outcomes of eligible studies were independently extracted by two reviewers (CW and CJ).

Methodological and Reporting Quality Assessment

Each of the studies was evaluated using the ‘‘Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials”[12] to access the HRQOL trial quality. The checklist consisted of 11 items grouped into four key categories related to the HRQOL assessment: conceptual, measurement, methodology and interpretation. This 11-item checklist was designed to have a dichotomous answer (yes / no): one mark for ‘yes’ (giving a score of 1) and zero mark for ‘no’ (giving a score of 0). Each study was classified into one of the following three descriptive categories: “very limited” (with a score between 0 and 4), “limited” (with a score between 5 and 7) and “probably robust” (with a score between 8 and 11 and with three mandatory items of the checklist: baseline compliance, psychometric properties reported and missing data documented)[12]. As a result, studies having a score of 8 or above but not possessing those three mandatory items were regarded as “limited”. This checklist provided a general guideline for addressing the basic and essential issues a study should possess in order to have convincing and significant outcomes in the assessment methodology.

Besides, authors (CW and CJ) reached consensus on which further quality assessment should be evaluated in comparative studies reporting HRQOL differences between CRC

patients and controls. A further quality assessment consists of five pre-defined quality assessment criteria: 1) Comparing with their populations: CRC and control groups should come from an identical source population; 2) Contemporary comparisons: both groups should be enrolled during the same time period of within 5 years; 3) General population: as source of the comparison group; 4) Matched comparison group: minimizing confounding factors that could introduce bias of differences between the CRC and control groups; 5) Reporting and presentation of results: results of CRC group, comparison group and the difference between groups should be reported and the statistical significance between groups should be tested. For each study, every single quality criterion was rated as "Yes" or "No" if the criterion was met or not respectively.

Results

Figure 1 lists the process of literature identification, screening for eligibility, and selection of studies during the literature search presented in a Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) flow diagram. The literature search was completed in June 2014 and identified a total of 7,553 potentially relevant studies (PubMed: 1,349; Web of Science: 2,318; MEDLINE: 1,735; and Embase: 2,151) that met the searching criteria in four bibliographic databases. Abstract screening removed the duplicated articles (n = 3,332), non-original articles (n = 1,439), and articles not related to CRC patients (n = 1,346) and non-comparative studies (n = 1,401). The full-text content of 35 studies was reviewed for eligibility. To exclude ineligible studies due to the use of non-standardized HRQOL instrument (n=1) and no general population/healthy controls for comparisons (n=3), the full-text articles of all eligible studies (n=31) were included. The earliest comparative study

that assessed the difference in HRQOL between CRC patients and controls was published in 2003.

Demographics and trial design characteristics

Thirty-one included studies compared various aspects of HRQOL in patients with CRC with general population or healthy controls. Three-fourth (n=23; 74.2%) of studies[13-36] originated from European countries, particularly in Netherland (n=6; 19.4%)[16, 23, 28-31], Germany (n=4; 12.9%)[13, 14, 20, 21], Italy (n=4; 12.9%)[17, 24, 25, 27] and Scandinavia (n=3; 9.7%). The remainders were originated from the US (n=5; 16.1%)[6, 22, 37-39], Australia (n=2; 6.5%)[40, 41], and Asia (n=1; 3.2%)[42]. Twenty-eight studies (90.3%) collected a sample of CRC patients and compared them to normative data published elsewhere[13-31, 33-42], while the remaining minority recruited a group of CRC patients and a group of healthy control subjects within the same studies. A summary of these 31 comparative studies is presented in Table 1.

Table 2 shows the general characteristics and available normative data of the standardized validated HRQOL instruments identified in comparative studies. The European Organisation for Research and Treatment of Cancer Quality-of-Life Questionnaire Core 30 (EORTC QLQ-C30) was the most extensively used HRQOL instrument, which was used in over half (n=16; 51.6%) of the studies [13-15, 18-22, 24-26, 31, 32, 34-36]. The European Organisation for Research and Treatment of Cancer Colorectal Cancer Specific Quality-of-Life Questionnaire module (EORTC QLQ-CR38) was also applied in 13 (43.3%) studies [14, 16, 19, 22-26, 29-32, 35]. Six non-cancer-specific HRQOL instruments such as EuroQoL 5-dimension (EQ-5D), SF-36, SF-12, SF-6D, Psychological General Well-Being Index (PGWBI), Hospital Anxiety and Depression Scale (HADS) were identified for HRQOL comparisons in

comparative studies. Aforementioned eight instruments were standardized and validated HRQOL instruments, in which those translations are available in many languages. Two studies[34, 42] reported the HRQOL measured by Functional Assessment of Cancer Therapy-colorectal (FACT-C) instrument but normative values of FACT-C instrument were not available for comparisons.

Overview of HRQOL assessment methodology and methods of analysis

In Table 3, the results of the HRQOL assessment methodology and methods of analysis are summarized in four major categories: conceptual, measurement, methodology, and interpretation, according to the "Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials"[12]. Methodological limitations were identified in several aspects of the overall process of HRQOL assessment, particularly in terms of the conceptual, methodology, and interpretation.

Conceptual

In the conceptual criteria, our review figured out a poor reporting of details about the priori hypothesis and rationale for selecting a specific HRQOL measure and instrument administration. Only one (3.2%) out of 31 studies had a priori hypothesis stated [29] and 35.5% (n=11) of the studies provided a rationale for selecting the specific HRQOL instrument [13, 14, 17, 19, 20, 22, 26, 27, 32, 34, 40].

Measurement

In terms of measurement, although 38.7% (n=12) of the reports did not verify the cultural validity of the study [13, 14, 18-22, 24, 25, 32, 35, 36], all reports except one (96.7%) covered adequate HRQOL domains and reported psychometric properties relevant for comparison [6, 13-42]. Thus, there were no major limitations regarding to the aspect of measurement.

Methodology

In our review, 93.5% (n=29) of the included studies had the documentation of timing assessment for analysis [6, 13-17, 19-38, 40-42]. 87.1% (n=27) of the studies reported the instrument administration and baseline compliance [6, 13-42]. However, 64.5% (n=20) of the studies did not provide any details about HRQOL missing data during the course of analysis [6, 15, 17, 18, 23-25, 27-31, 33, 35, 37-42]. This led to the loss of clinically significant differences due to reduced number of observations.

Interpretation

All of the studies (100%) had adequate presentation and discussion of results in general[6, 13-42] but ten reports (32.3%) did not provide the clinical significance for analysis [17, 23, 31-34, 36, 37, 40, 42] which might limit the clear understanding of results and clinical relevance of HRQOL changes.

Overview of HRQOL trial quality and treatment recommendation on patient's HRQOL

Based on the minimum standard checklist score in Table 4, only one study (3.2%) was classified as very limited [35] and 71.0% (n = 22) were classified as limited [6, 15, 17, 18, 21, 23-25, 27-33, 36-42] while eight studies (25.8%) were classified as "probably robust" for clinical decision making[13, 14, 16, 19, 20, 22, 26, 34]. These studies demonstrated excellent examples for the implementation of HRQOL assessment of patients to thoroughly evaluate the overall effectiveness of treatment.

Further quality assessment

Table 5 depicts the further quality assessments of included studies. Based on five pre-defined quality assessment criteria, our study reviewed the poor reporting of result, comparison and matching of the patients groups. For the aspects of result and comparison, 61.3% (n=19) did not have the results of difference between groups reported [14, 16, 17, 21, 22, 25, 26, 28, 30-35, 37-41] and 36.7% (n=11) were not contemporary comparisons (no. of years > 5) [17, 20, 22, 24-27, 33-35, 39]. 38.7% (n=12) of the papers did not have the results of statistical significance tested [12-14, 21, 22, 24, 25, 34, 35, 37, 38, 41] while only 54.8% (n=17) were tested by univariate analysis [6, 15-20, 23, 26, 27, 29-33, 40, 42] and 6.5% (n=2) were tested by both univariate and multivariate analysis [6, 13-35, 37-42]. For the matching method, 58.1% (n=18) of studies did not give matching criteria and perform matching to identify comparison group [13, 14, 17, 19-22, 26, 27, 32-35, 37-41]. 16.7% (n=5) did not use population within the same countries as controls to compare [22, 24-26, 35] and 6.7% (n=2) did not adapt representative general population as source of comparison group[6, 32].

Normative Comparisons

Norms for cancer-specific HRQOL instrument QLQ-C30, which was used in about half (n=15; 48.4%) of the studies[13-15, 18-22, 24-26, 31, 32, 35, 36], have been obtained from the general population in German[43], Norway[44], Netherland[45], Austria[46], Sweden[47], and France[15]. Published normative data of QLQ-C30 were available in the general adult population after matching for age and sex [15, 18, 24, 25, 31, 36] and different aspects such as CRC patients after diagnosis [13-15, 20], CRC patients after surgery [21, 31, 32, 34, 36], rectal cancer patients following diagnosis [19, 26] and rectal cancer patients following surgery [22, 24, 25, 31, 35]. Moreover, reference data from the QLQ-C30 in a sample of 3000 adults from Norwegian general population provided comparison with rectal cancer patients during radiotherapy [19] or rectal cancer survivors [26]. Dutch normative data of QLQ-C30 in age- and sex-matched general population

were compared with random samples of Eindhoven Cancer Registry[31]. Besides, norms for another cancer-specific HRQOL instrument QLQ-CR38, which was also found in 5 (16.1%) studies[16, 23, 29-31], have been obtained from the general population in Netherland[45] for comparisons. Only one study recruited the healthy controls through random sampling in the general population[32] and tested the control group against the CRC patients after surgery.

Norms for generic HRQOL instruments are also available in multiple countries[45, 48-55]. General population norms for SF-36 Health Survey was compared with HRQOL of CRC patients in Netherland[23, 29, 30], France[15], Australia[41], Italy[17], Finland[33] and the US[6, 37, 38]. Comparisons with general population norms for SF-12 were utilized in UK[34], Australia[40] and Hong Kong[42]. The HADS score of CRC patients were compared to the UK[50] and Dutch[28] general population. The PGWBI, EQ-5D and SF-6D scores were used in one study in Italy [27], in UK and in the US[39], respectively.

Compared to the normative data from the general population of Italy [24, 25], Germany [13, 14, 21], Norway [19], Austria [18] and Poland [35], colorectal cancer patients had worse scores in most of functioning and symptom scales measured by QLQ-C30. Cancer survivors had significantly lower physical component summary (PCS) score compared to the Australian general population [41] and the United Kingdom population with an interval of age 65-74 [34]. Inconsistent results were observed in the US population when the generic and cancer-specific HRQOL of general population were similar to that of CRC patients [22, 37, 38] but SF-6D score norms ranging from 0.76 to 0.80 were higher than that with 0.69 for permanent stoma and 0.73 for anastomses among CRC survivors [39]. Conversely, older CRC survivors reported better compared to the general population in Germany [20] and Finland [33] but younger survivors

conversely reported worse. The HRQOL of rectal cancer survivors who have completed treatment regimen for more than two years or ten years were higher than that of the general population from France[26] and Netherlands[29, 30], respectively. In another French population-based study, cancer survivors reported worse social functioning and more diarrhea symptoms at five years after diagnosis compared to the healthy control group [15]. Healthy controls had significantly better results in functional and symptom scales of QLQ-C30 than patients with CRC in Bosnia[32] and Austria[18].

For QLQ-CR38, there were almost no comparisons in those studies between the CRC group and control group except one study [32]. CRC patients had worse scores in most of scales measured by QLQ-CR38 compared to healthy controls. For SF-12, CRC patients had lower PCS but higher mental component summary (MCS) scores compared to the norms of their respective countries in Australia, UK and Hong Kong[34, 40, 42]. For SF-36, Italian patients with colorectal cancer had lower HRQOL in reference to bodily pain, social functioning and general health measured than that of general population[17]. For Australian rectal cancer patients, they had lower PCS but similar MCS to that of general population [41] while there were no big difference between the CRC survivors group and control group for the American female patients [37, 38]. For PGWBI, population-based reference data was collected in the Italian general population as control group to compare with rectal cancer survivors who reported better scores on all PGWBI scales except for self-control scale [27].

Discussions

In this systematic review, we summarized and appraised the methodological and reporting quality of 31 studies that compared the HRQOL between CRC patients and controls,

which provided information on the additional impact of CRC on HRQOL. The results have shown that there were only 8 studies (25.8%) considered as "probably robust" regarding the methodological and reporting quality of HRQOL comparisons, hampering the informing clinical practice and decision making.

This review detected several drawbacks of current studies reporting HRQOL comparisons between CRC and control groups. First, there is a disparity of the HRQOL instrument used. High-quality studies were more likely to measure HRQOL using the EORTC QLQ-C30 instrument in conjunction with QLQ-CR38 instrument while several studies only used generic HRQOL instruments to compare the HRQOL between the CRC and control groups (references only use generic HRQOL). It is more informative to combine both generic and specific HRQOL instrument in comparing the HRQOL between the CRC and control groups because generic instruments of HRQOL can be used to compare HRQOL over a broad spectrum of diseases, as well as general population, and was more responsive to detect changes in social domain than colorectal-specific HRQOL instrument[56]. However, only 5 studies used both generic and specific HRQOL instruments based on the results of our systematic review. Secondly, almost 40% of the included studies used HRQOL instrument that has not been culturally verified. It is important to choose a well-tested HRQOL measure in certain culture, as it is culture-dependent[57].

This review underlined the importance of establishing an appropriate control group for fair comparison. Of the 31 included studies, only three studies recruited the CRC patients and healthy controls within the same studies[15, 18, 32]. Some studies even compare reference data from different populations, which may not accurately reflect the additional impact of CRC on

HRQOL compared to the controls in the same country. Since HRQOL is always dependent on age and sex, comparison between CRC patients and the control group should match with age and sex, or adjust for multiple covariates using propensity score. However, over 60% of the studies did not performed matching strategy to identify the controls. Another issue of the control group is relevant to the non-contemporary comparison, which means that the difference of the recruitment period between CRC group and the control group is over 5 years.

Statistical significance is useful in interpreting the data to be accounted for fluctuations by chance, and thus does not necessarily imply the clinical significance. A difference that is statistically different may have little or even no importance in the realm of health care and health decision-making. In this systematic review, ten comparative studies (32.3%) did not provide the clinical significance for analysis [17, 23, 31-34, 36, 37, 40, 42], which may hamper the clinical meaningfulness based on the results. Clinical significance of HRQOL scores was determined by two main approaches. For those studies interpreting whether changes were considered as clinical significance, a half-standard deviation approach[58], corresponding to Cohen's medium effect size, was adopted[6, 16, 28-30] for detecting clinically important difference of HRQOL scores. In comparative studies[13-15, 18-20, 22, 24, 25] administering EORTC instruments, scores difference of at least 10 absolute points is interpreted as a clinically important difference according to Osoba et al[59].

Priori hypothesis and rational for selection a HRQOL instrument are lacking in most of the included studies. As the priori hypothesis is the key prerequisite for deciding which HRQOL instrument to be used, the lack of such information may lead to spurious positive results

because of the multiple tests in comparing different HRQOL domains between the CRC and control groups.

Limitations of this review should be noted. First, methodological quality assessment relied on the information reported in published articles which may be shortened subject to the editorial and reviewers' request. Results of published articles may be partially reported. Further eta-analysis of differences between the two groups was not available based on HRQOL point estimate reported in published articles.

Conclusions

In conclusion, this review showed that one-fourth of comparative studies generally achieved high-quality in reporting characteristics and methodological details. HRQOL is increasingly used to complement outcomes of CRC patients, but our systematic review noted that only 8 out of 31 studies met the methodological criteria as probably robust for clinical decision making according to the Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials. Researchers should pay careful attention to the HRQOL instrument standardization with a priori hypothesis, and to choose a comparable control group with similar culture background recruited at a similar time point. Future studies investigating the impact of CRC on HRQOL are encouraged to undertake HRQOL measurement and adhere with methodological checklist and further pre-defined assessment criteria.

Reference

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer*. 2010;127:2893-2917.
2. Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin*. 2014;64:104-117.
3. Byrne C, Griffin A, Blazeby J, Conroy T, Efficace F. Health-related quality of life as a valid outcome in the treatment of advanced colorectal cancer. *Eur J Surg Oncol*. 2007;33:S95-S104.
4. Basch E, Abernethy AP, Mullins CD, et al. Recommendations for incorporating patient-reported outcomes into clinical comparative effectiveness research in adult oncology. *J Clin Oncol*. 2012;30:4249-4255.
5. Baker F, Haffer SC, Denniston M. Health-related quality of life of cancer and noncancer patients in Medicare managed care. *Cancer*. 2003;97:674-681.
6. Reeve BB, Potosky AL, Smith AW, et al. Impact of cancer on health-related quality of life of older Americans. *J Natl Cancer Inst*. 2009;101:860-868.
7. Jansen L, Koch L, Brenner H, Arndt V. Quality of life among long-term (≥5 years) colorectal cancer survivors – systematic review. *Eur J Cancer*. 2010;46:2879-2888.
8. Harji DP, Griffiths B, Velikova G, Sagar PM, Brown J. Systematic review of health-related quality of life issues in locally recurrent rectal cancer. *J Surg Oncol*. 2015;111:431-438.
9. Efficace F, Bottomley A, Osoba D, et al. Beyond the development of health-related quality-of-life (HRQOL) Measures: A checklist for evaluating HRQOL outcomes in cancer clinical trials—Does HRQOL evaluation in prostate cancer research inform clinical decision making? *J Clin Oncol*. 2003;21:3502-3511.
10. Efficace F, Bottomley A, Vanvoorden V, Blazeby JM. Methodological issues in assessing health-related quality of life of colorectal cancer patients in randomised controlled trials. *Eur J Cancer*. 2004;40:187-197.
11. Wong CKH, Chen J, Yu CLY, Sham M, Lam CLK. Systematic review recommends the European Organization for Research and Treatment of Cancer colorectal cancer–specific module for measuring quality of life in colorectal cancer patients. *J Clin Epidemiol*. 2015;68:266-278.
12. Efficace F, Horneber M, Lejeune S, et al. Methodological quality of patient-reported outcome research was low in complementary and alternative medicine in oncology. *J Clin Epidemiol*. 2006;59:1257.e1251-1257.e1214.

13. Arndt V, Merx H, Stegmaier C, Ziegler H, Brenner H. Quality of Life in patients with colorectal cancer 1 year after diagnosis compared with the general population: a population-based study. *J Clin Oncol*. 2004;22:4829-4836.
14. Arndt V, Merx H, Stegmaier C, Ziegler H, Brenner H. Restrictions in quality of life in colorectal cancer patients over three years after diagnosis: a population based study. *Eur J Cancer*. 2006;42:1848-1857.
15. Caravati-Jouvenceaux A, Launoy G, Klein D, et al. Health-related quality of life among long-term survivors of colorectal cancer: A population-based study. *Oncologist*. 2011;16:1626-1636.
16. Den Oudsten BL, Traa MJ, Thong MS, et al. Higher prevalence of sexual dysfunction in colon and rectal cancer survivors compared with the normative population: a population-based study. *Eur J Cancer*. 2012;48:3161-3170.
17. Domati F, Rossi G, Benatti P, Roncucci L, Cirilli C, de Leon MP. Long-term survey of patients with curable colorectal cancer with specific reference to the quality of life. *Intern Emerg Med*. 2011;6:529-535.
18. Giesinger J, Kemmler G, Mueller V, et al. Are gender-associated differences in quality of life in colorectal cancer patients disease-specific? *Qual Life Res*. 2009;18:547-555.
19. Guren MG, Dueland S, Skovlund E, Fossa SD, Poulsen JP, Tveit KM. Quality of life during radiotherapy for rectal cancer. *Eur J Cancer*. 2003;39:587-594.
20. Jansen L, Herrmann A, Stegmaier C, Singer S, Brenner H, Arndt V. Health-related quality of life during the 10 years after diagnosis of colorectal cancer: a population-based study. *J Clin Oncol*. 2011;29:3263-3269.
21. Kopp I, Bauhofer A, Koller M. Understanding quality of life in patients with colorectal cancer: comparison of data from a randomised controlled trial, a population based cohort study and the norm reference population. *Inflamm Res*. 2004;53 S130-135.
22. 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? *Ann Surg Oncol*. 2007;14:1727-1734.
23. Orsini RG, Thong MS, van de Poll-Franse LV, et al. Quality of life of older rectal cancer patients is not impaired by a permanent stoma. *Eur J Surg Oncol*. 2013;39:164-170.
24. Pucciarelli S, Del Bianco P, Efficace F, et al. Health-related quality of life, faecal continence and bowel function in rectal cancer patients after chemoradiotherapy followed by radical surgery. *Support Care Cancer*. 2010;18:601-608.
25. Pucciarelli S, Del Bianco P, Toppan P, et al. Health-related quality of life outcomes in disease-free survivors of mid-low rectal cancer after curative surgery. *Ann Surg Oncol*. 2008;15:1846-1854.
26. Rauch P, Miny J, Conroy T, Neyton L, Guillemin F. Quality of life among disease-free survivors of rectal cancer. *J Clin Oncol*. 2004;22:354-360.

27. Serpentini S, Del Bianco P, Alducci E, et al. Psychological well-being outcomes in disease-free survivors of mid-low rectal cancer following curative surgery. *Psychooncology*. 2011;20:706-714.
28. Thong MS, Mols F, Wang XS, Lemmens VE, Smilde TJ, van de Poll-Franse LV. Quantifying fatigue in (long-term) colorectal cancer survivors: a study from the population-based patient reported outcomes following initial treatment and long term evaluation of survivorship registry. *Eur J Cancer*. 2013;49:1957-1966.
29. Thong MSY, Mols F, Lemmens VEPP, Creemers G-J, Slooter GD, van de Poll-Franse LV. Impact of chemotherapy on health status and symptom burden of colon cancer survivors: a population-based study. *Eur J Cancer*. 2011;47:1798-1807.
30. Thong MSY, Mols F, Lemmens VEPP, et al. Impact of preoperative radiotherapy on general and disease-specific health status of rectal cancer survivors: a population-based study. *Int J Cancer*. 2011;81:e49-58.
31. Traa MJ, Orsini RG, Den Oudsten BL, et al. Measuring the health-related quality of life and sexual functioning of patients with rectal cancer: does type of treatment matter? *Int J Cancer*. 2014;134:979-987.
32. Trninic Z, Vidacak A, Vrhovac J, Petrov B, Setka V. Quality of life after colorectal cancer surgery in patients from University Clinical Hospital Mostar, Bosnia and Herzegovina. *Coll Antropol*. 2009;33 1-5.
33. Vironen JH, Kairaluoma M, Aalto AM, Kellokumpu IH. Impact of functional results on quality of life after rectal cancer surgery. *Dis Colon Rectum*. 2006;49:568-578.
34. 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;49:1692-1702.
35. Zajac O, Spychala A, Murawa D, Wasiewicz J, Foltyn P, Polom K. Quality of life assessment in patients with a stoma due to rectal cancer. *Rep Pract Oncol Radiother*. 2008;13:130-134.
36. Pollack J, Holm T, Cedermark B, et al. Late adverse effects of short-course preoperative radiotherapy in rectal cancer. *Br J Surg*. 2006;93:1519-1525.
37. Sapp AL, Trentham-Dietz A, Newcomb PA, Hampton JM, Moinpour CM, Remington PL. Social networks and quality of life among female long-term colorectal cancer survivors. *Cancer*. 2003;98:1749-1758.
38. Trentham-Dietz A, Remington PL, Moinpour CM, Hampton JM, Sapp AL, Newcomb PA. Health-related quality of life in female long-term colorectal cancer survivors. *Oncologist*. 2003;8:342-349.
39. Hornbrook MC, Wendel CS, Coons SJ, et al. Complications among colorectal cancer survivors: SF-6D preference-weighted quality of life scores. *Med Care*. 2011;49:321-326.
40. Gall CA, Weller D, Esterman A, et al. Patient satisfaction and health-related quality of life after treatment for colon cancer. *Dis Colon Rectum*. 2007;50:801-809.

41. Austin KKS, Young JM, Solomon MJ. Quality of life of survivors after pelvic exenteration for rectal cancer. *Dis Colon Rectum*. 2010;53:1121-1126.
42. Wong CK, Lam CL, Poon JT, Kwong DL. Clinical correlates of health preference and generic health-related quality of life in patients with colorectal neoplasms. *PLoS One*. 2013;8:e58341.
43. Schwarz R, Hinz A. Reference data for the quality of life questionnaire EORTC QLQ-C30 in the general German population. *Eur J Cancer*. 2001;37:1345-1351.
44. Hjermstad MJ, Fayers PM, Bjordal K, Kaasa S. Health-related quality of life in the general Norwegian population assessed by the European Organization for Research and Treatment of Cancer Core Quality-of-Life Questionnaire: the QLQ=C30 (+ 3). *J Clin Oncol*. 1998;16:1188-1196.
45. van de Poll-Franse LV, Mols F, Gundy CM, et al. Normative data for the EORTC QLQ-C30 and EORTC-sexuality items in the general Dutch population. *Eur J Cancer*. 2011;47:667-675.
46. Holzner B, Kemmler G, Cella D, et al. Normative data for functional assessment of cancer therapy General scale and its use for the interpretation of quality of life scores in cancer survivors. *Acta Oncol*. 2004;43:153-160.
47. Michelson H, Bolund C, Nilsson B, Brandberg Y. Health-related quality of life measured by the EORTC QLQ-C30: Reference values from a large sample of the Swedish population. *Acta Oncol*. 2000;39:477-484.
48. Crawford JR, Henry JD, Crombie C, Taylor EP. Normative data for the HADS from a large non-clinical sample. *Br J Clin Psychol*. 2001;40:429-434.
49. Apolone G, Mosconi P. The Italian SF-36 Health Survey: translation, validation and norming. *J Clin Epidemiol*. 1998;51:1025-1036.
50. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. *BMJ*. 1998;316:736-741.
51. Ware JE. SF-36 Health Survey. Manual and interpretation guide. Second ed. The Health Institute, New England Medical Centre. Boston, MA.: Nimrod Press, 1997.
52. Hawthorne G, Osborne RH, Taylor A, Sansoni J. The SF36 Version 2: critical analyses of population weights, scoring algorithms and population norms. *Qual Life Res*. 2007;16:661-673.
53. Aro AR, Teperi J, Aro AR. Rand-36 terveyteen liittyvän elämänlaadun mittarina: Mittarin luotettavuus ja suomalaiset väestöarvot. 1999.
54. Sanderson K, Andrews G. The SF-12 in the Australian population: cross-validation of item selection. *Aust N Z J Public Health*. 2002;26:343-345.
55. Lam CLK, Wong CK, Lam ETP, Lo YYC, Huang WW. Population norm of Chinese (HK) SF-12 health survey-version 2 of Chinese adults in Hong Kong. *Hong Kong Practitioner*. 2010;32:77-86.

56. Wong CK, Lam CL, Law W-L, et al. Condition-specific measure was more responsive than generic measure in colorectal cancer: all but social domains. *J Clin Epidemiol*. 2013;66:557-565.
57. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46:1417-1432.
58. Norman GR, Sloan JA, Wyrwich KW. Interpretation of changes in health-related quality of life: The remarkable universality of half a standard deviation. *Med Care*. 2003;41:582-592.
59. Osoba D, Rodrigues G, Myles J, Zee B, Pater J. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol*. 1998;16:139-144.

Running title: Quality of CRC HRQOL comparative studies

Running title: Quality of CRC HRQOL comparative studies

Table 1. Study Characteristics of 31 Eligible Studies

First Author	Year of Publication	Country of origin	Study design of CRC data	Target Population: CRC	Target Population: Trajectory Stage of CRC	Target Population: Compared with	Response rate	Sample Size: CRC	Sample Size: non-CRC	Age (year): CRC	Age (year): Non-CRC	Male (%): CRC	Male (%): Non-CRC	HRQOL instrument
Arndt 2004[13]	2004	Germany	Prospective	CRC	1 Year After Diagnosis	German general population[43]	81.7%	309	2028	Mean 65.1 ± 9.4 (18–80)	Mean 66.0 ± 11.6 (16–92)	56.3%	43.8%	EORTC QLQ-C30
Arndt 2006[14]	2006	Germany	Prospective	CRC	3 Year After Diagnosis	German general population[43]	92.1%	222	2028	Mean 66.0 ± 9.2	Mean 66.0 ± 11.6 (16–92)	52%	43.8%	EORTC QLQ-C30 EORTC QLQ-CR38
Jansen[20]	2011	Germany	Prospective	CRC	10 Years After Diagnosis	German general population[43]	60%	117	2028	Mean 62.6 ± 8.9	Mean 66.0 ± 11.6 (16–92)	46%	43.8%	EORTC QLQ-C30
Kopp[21]	2004	Germany	RCT and Cohort studies	CRC	After Surgery	German general population[43]	NR	325	193	Mean 68.6 (33–92)	Range 60–69	62.8%	100%	EORTC QLQ-C30
Rauch[26]	2004	France	Cross-sectional	Rectum	Disease-free survivors and complete remission more than 2 years after diagnosis	German general population[43] Norwegian general population[44]	78.1%	121	3993	Median 64 (43–91)	Norwegian: Mean 47.4 (19–93) German: Mean 66.0 ± 11.6 (16–92)	64.5%	47.7%	EORTC QLQ-C30 EORTC QLQ-CR38
Neuman[22]	2007	US	Retrospective	Rectum	After sphincter-preserving surgery	German general population[43]	67.4%	123	2028	Median 63.0 (35–87)	Mean 66.0 ± 11.6 (16–92)	67.5%	43.8%	EORTC QLQ-C30 EORTC QLQ-CR38
Pucciarel	2008	Italy	Retrospective	Rectum	After Surgery	Age- and	80.7%	117	117	Median 65	NR	63.2%	NR	EORTC

				Target Population				Sample Size		Age (year)		Male (%)		
First Author	**Year of Publication**	**Country of origin**	**Study design of CRC data**	**CRC**	**Trajectory Stage of CRC**	**Compared with**	**Response rate**	**CRC**	**non-CRC**	**CRC**	**Non-CRC**	**CRC**	**Non-CRC**	**HRQOL instrument**
li 2008[25]						sex-matched German general population[43]				(39–92)				QLQ-C30 EORTC QLQ-CR38
Zajac[35]	2008	Poland	Cross-sectional	Rectum	After Surgery	German general population[43]	NR	50	2028	Mean 62.1 (38 to 80)	Mean 66.0 ± 11.6 (16–92)	56.0%	43.8%	EORTC QLQ-C30 EORTC QLQ-CR38
Pucciarelli 2010[24]	2010	Italy	Cross-sectional	Rectum	After chemo-radiotherapy followed by radical surgery	Age- and sex-matched German general population[43]	80.2%	81	81	Median 62 (33–81)	NR	58.0%	NR	EORTC QLQ-C30 EORTC QLQ-CR38
Guren[19]	2003	Norway	Prospective	Rectum	During Radiotherapy	Norwegian general population[44]	75.0%	42	1965	Median 67 (38–78)	Mean 47.4 (19–93)	59.5%	51.7%	EORTC QLQ-C30 EORTC QLQ-CR38
Thong 2011a[30]	2011	Netherlands	Random sample survey on Eindhoven Cancer Registry	Rectum	Survivors	Age- and sex-matched Dutch general population[45]	62.2%	340	1731	Mean 68.2 ± 9.6	Mean 53 ± 16	66.2%	54.0%	EORTC QLQ-CR38 SF-36

				Target Population				Sample Size		Age (year)		Male (%)		
First Author	**Year of Publication**	**Country of origin**	**Study design of CRC data**	**CRC**	**Trajectory Stage of CRC**	**Compared with**	**Response rate**	**CRC**	**non-CRC**	**CRC**	**Non-CRC**	**CRC**	**Non-CRC**	**HRQOL instrument**
Thong 2011b[29]	2011	Netherlands	Random sample survey on Eindhoven Cancer Registry	Colon	Survivors	Age- and sex-matched Dutch general population[45]	74.7%	848	1731	Mean 69.4 ± 9.6	Mean 53 ± 16	43.8%	54.0%	EORTC QLQ-CR38 SF-36
Austin[41]	2010	Australia	Cross-sectional	Rectum	Survivors After Pelvic Exenteration	Australian general population[52]	84.1%	37	3014	Median 62 (31–85)	Mean 45.29 ± 18.69	43.2%	49.1%	SF-36
Domati[17]	2011	Italy	Retrospective	CRC	After Surgery (5 years after the diagnosis)	Italian general population[49]	38.8%	220	NR	Mean 66.6 (43–81)	NR	57.3%	NR	SF-36
Vironen[33]	2006	Finland	Cross-sectional	Rectum	After Surgery	Finnish general population[53]	87.2%	82	1440	Mean 68	Range 40-79	63.4%	43.3%	SF-36
Sapp[37]	2003	US	Cross-sectional	Female CRC	Survivors	US general woman population[51]	94.9%	209	413	Mean 72 (43–85)	≥65	0.0%	0.0%	SF-36
Trentham-Dietz[38]	2003	US	Cross-sectional	Female CRC	Survivors	US general woman population[51]	94.9%	209	413	Mean 72 (43–85)	≥65	0.0%	0.0%	SF-36
Wilson[34]	2006	UK	Prospective	CRC	After Surgery	UK general population[50]	95.7%	201	NR	Mean 68.2 (36–91)	Range 65-74	73.1%	NR	SF-12 EQ-5D EORTC QLQ-C30 FACT-C

				Target Population				Sample Size		Age (year)		Male (%)		
First Author	**Year of Publication**	**Country of origin**	**Study design of CRC data**	**CRC**	**Trajectory Stage of CRC**	**Compared with**	**Response rate**	**CRC**	**non-CRC**	**CRC**	**Non-CRC**	**CRC**	**Non-CRC**	**HRQOL instrument**
Serpentini[27]	2011	Italy	Cross-sectional	Rectum	Survivors	Italian general population[49]	80.7%	117	1129	Median 65	15-24 (18.5%) 25-44 (35.3%) 45-64 (28.6%) ≥65 (17.6%)	63.2%	48.1%	PGWBI
Hornbrook[39]	2011	US	Cross-sectional	CRC	Survivors	US general population[51]	51.9%	679	NR	Ostomate: Mean 72 ± 10 Nonostomate: Mean 71 ± 11	NR	58.9%	NR	SF-6D
Giesinger[18]	2009	Austria	Cross-sectional	CRC	Unknown	Age- and sex-matched Austrian general population[46]	NR	206	206	Mean 64.8 ± 11.5 (33–88)	Mean 64.9 ± 11.6	52.9%	52.9%	EORTC QLQ-C30
Trninic[32]	2009	Bosnia and Herzegovina	Cross-sectional	CRC	After Surgery	Healthy population	76.3%	67	30	With colostomy: Mean 64 ± 12.9 Without colostomy : Mean 61 ±	Mean 60 ± 12.2	51.7%	51.7%	EORTC QLQ-C30 EORTC QLQ-CR38

First Author	Year of Publication	Country of origin	Study design of CRC data	Target Population: CRC	Target Population: Trajectory Stage of CRC	Target Population: Compared with	Response rate	Sample Size: CRC	Sample Size: non-CRC	Age (year): CRC	Age (year): Non-CRC	Male (%): CRC	Male (%): Non-CRC	HRQOL instrument
										12.7				
Caravati-Jouvenceaux[15]	2011	France	Random sample survey from three tumor registries in France	CRC	5, 10, and 15 Years After Diagnosis	Age-, sex- and residence area-matched French general population	37.2%	542	1181	Mean 70.8	Mean 70.2	56.6%	50.9%	EORTC QLQ-C30 SF-36
Gall[40]	2007	Australia	Prospective	Colon	After recovery from treatment	Australian general population[54] for SF-12 UK general population for HADS[48]	74.3%	338	NR	<60 (12.7%) 60-69 (22.8%) 70-79 (43.8%) ≥80 (20.7%)	NR	56.0%	NR	SF-12 HADS
Reeve[6]	2009	US	Random sample from Medicare Health Outcomes Survey	CRC	first cancer diagnosis occurred between their baseline and follow-up MHOS assessments	Propensity matched control subjects without cancer	NR	240	7160	>65	Mean 73.81 ± 6.04	NR	55.5%	SF-36

				Target Population				Sample Size		Age (year)		Male (%)		
First Author	**Year of Publication**	**Country of origin**	**Study design of CRC data**	**CRC**	**Trajectory Stage of CRC**	**Compared with**	**Response rate**	**CRC**	**non-CRC**	**CRC**	**Non-CRC**	**CRC**	**Non-CRC**	**HRQOL instrument**
Den Oudsten[16]	2012	Netherlands	Random sample survey on Eindhoven Cancer Registry	CRC	Survivors	Age- and sex-matched Dutch general population[45]	81.5%	1371	400	Mean 70 ± 10	Mean 69 ± 10	55.8%	56.0%	EORTC QLQ-CR38
Thong 2013[28]	2013	Netherlands	Random sample survey on Eindhoven Cancer Registry	CRC	Survivors	Age- and sex-matched Dutch general population[45]	79.8%	3739	338	Mean 70 ± 10	Mean 68 ± 11	36.4%	55.6%	HADS
Orsini[23]	2013	Netherlands	Random sample survey on Eindhoven Cancer Registry	Rectum	Survivors	Age- and sex-matched Dutch general population[45]	91.7%	143	1613	Mean 64.7 ± 10.1	NR	62.2%	NR	SF-36 EORTC QLQ-CR38
Traa[31]	2014	Netherlands	Cross-sectional	Rectum	After Surgery	Age- and sex-matched Dutch general population[45]	85%	439	350	Mean 66.2 ± 9.8	Mean 66.4 ± 10.4	59.0%	57.1%	EORTC QLQ-C30 EORTC QLQ-CR38
Wong[42]	2013	Hong Kong, China	Cross-sectional	CRC	Survivors	Age- and sex-matched Hong Kong general population[55]	79.4%	381	515	Mean 64.3 ± 11.0	NR	54.6%	NR	SF-12 FACT-C

				Target Population				Sample Size		Age (year)		Male (%)		
First Author	**Year of Publication**	**Country of origin**	**Study design of CRC data**	**CRC**	**Trajectory Stage of CRC**	**Compared with**	**Response rate**	**CRC**	**non-CRC**	**CRC**	**Non-CRC**	**CRC**	**Non-CRC**	**HRQOL instrument**
Pollack[36]	2006	Sweden	Cross-sectional	Rectum	After Surgery	Age-matched sample of the Swedish population[47]	55.2%	139	NR	mean 74	NR	54.0%	NR	EORTC QLQ-C30

Note: NR=Not reported; CRC=colorectal cancer; HRQOL=health-related quality of life; EORTC= European Organisation for Research and Treatment of Cancer; QLQ-C30=quality of life questionnaire core 30 module; QLQ-CR38= colorectal cancer specific quality of life questionnaire module; SF-12= 12-item Short term Health Survey; SF-36= 36-item Short term Health Survey; SF-6D= Short term 6-dimension Health Survey; EQ-5D= EuroQoL 5-dimension; FACT-C= Functional Assessment of Cancer Therapy-colorectal; HADS= Hospital Anxiety and Depression Scale; PGWBI= Psychological General Well-Being Index

Running title: Quality of CRC HRQOL comparative studies

Table 2. General characteristics and available normative data of the standardized validated HRQOL instruments identified in 31 comparative studies

Instrument	Target Disease Population	Items	Subscales/Domains Score	Single Item Score	Response Options	Original language	Available normative data
Cancer-specific							
EORTC QLQ-C30	Cancer	30	9	6	4-point (28 items) 7-point (2 global items)	English	German[43], Norway[44], Netherland[45], Austria[46], Sweden[47], France[15]
EORTC QLQ-CR38	CRC	38	9	3	4-point	English	Netherland[45]
Non-cancer-specific							
HADS	Anxiety Disorders / Depression	14	2	0	4-point	English for UK	UK[48], Netherland[45]
Generic							
PGWBI	General	22	6	0	6-point	English for the US	Italy[49]
EQ-5D	General	5	5	0	3-point	English	UK[50]
SF-36	General	5	5	0	3-point (10 items) 5-point (25 items) 6-point (1 item)	English for the US	US[51], Australia[52], Finland[53], Netherland[45], France
SF-12	General	12	8	0	3-point (2 items) 5-point (10 items)	English for the US	UK[50], Australia[54], Hong Kong[55]
SF-6D	General	6	6	0	4-point (1 item) 5-point (3 items) 6-point (2 items)	English	US[51]

Note:

EORTC=European Organization for Research and Treatment of Cancer; QLQ=Quality-of-Life Questionnaire; SF-12= 12-item Short term Health Survey; SF-36= 36-item Short term Health Survey; SF-6D= Short term 6-dimension Health Survey; EQ-5D= EuroQoL 5-dimension; HADS= Hospital Anxiety and Depression Scale; PGWBI= Psychological General Well-Being Index

Table3

Running title: Quality of CRC HRQOL comparative studies

Table 3. Level of reporting of minimum standard checklist for evaluation of HRQOL outcomes in cancer clinical trials

HRQOL Issue	Criteria for evaluating items	No. (out of 31)	%
Conceptual			
A priori hypothesis stated	Assessed if authors had a predefined PRO endpoint and/or stated expected changes due to the specific treatment	1	3.2%
Rationale for instrument reported	Assessed if authors gave a rationale for using a specific PRO measure	11	35.5%
Measurement			
Psychometric properties reported	Assessed if a previously validated measure was used or psychometric properties were reported or referenced in the paper	30	96.8%
Cultural validity verified	Assessed if the measure was validated for the specific study population	19	61.3%
Adequacy of domains covered	Assessed if the measure covered, at least, the main PRO dimensions relevant for a generic cancer population and/or according to the specific research question	30	96.8%
Methodology			
Instrument administration reported	Assessed if authors specified who and/or in which clinical setting the PRO instrument was administered	27	87.1%
Baseline compliance reported	Assessed if authors reported the number of patients providing a PRO assessment before the start of treatment	27	87.1%
Timing of assessments documented	Assessed if authors specified the PRO timing of assessment during the trial	29	93.5%
Missing data documented	Assessed if authors gave some details on PRO missing data during the trial	11	35.5%
Interpretation			
Clinical significance addressed	This refers to the discussion of PRO data being clinically significant from a patient's perspective and not simply statistically significant	21	67.7%
Presentation of results in general	Assessed if authors discussed the PRO outcomes giving any comments regardless of the results (either expected or not)	31	100.0%

Note: HRQOL=health-related quality of life; PRO=patient-reported outcome

Running title: Quality of CRC HRQOL comparative studies

Table 4. Minimum Standard Checklist for Evaluating HRQOL Outcomes in Cancer Clinical Trials

First Author	Conceptual		Measurement			Methodology				Interpretation		Check list Score	Expected methodological quality
	A priori hypothesis stated (Yes/No/NA)	Rationale for instrument reported (Yes/No)	Psychometric properties reported (Yes/No)	Cultural validity verified (Yes/No/NA)	Adequacy of domains covered (Yes/No)	Instrument administration reported (Yes/No)	Baseline compliance reported (Yes/No)	Timing of assessments documented (Yes/No)	Missing data documented (Yes/No)	Clinical significance addressed (Yes/No)	Presentation of results in general (Yes/No)		
Arndt 2004[13]	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9 / 11	Probably robust
Arndt 2006[14]	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9 / 11	Probably robust
Jansen[20]	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9 / 11	Probably robust
Kopp[21]	No	No	Yes	No	Yes	No	No	Yes	Yes	Yes	Yes	6 / 11	Limited
Rauch[26]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	10 / 11	Probably robust
Neuman[22]	No	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9 / 11	Probably robust
Pucciarelli 2008[25]	No	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	7 / 11	Limited
Zajac[35]	No	No	No	No	Yes	No	No	Yes	No	Yes	Yes	4 / 11	Very limited
Pucciarelli 2010[24]	No	No	Yes	No	Yes	Yes	Yes	Yes	No	Yes	Yes	7 / 11	Limited
Guren[19]	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	Yes	Yes	8 / 11	Probably robust
Thong 2011a[30]	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	9 / 11	Limited
Thong 2011b[29]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8 / 11	Limited
Austin[41]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8 / 11	Limited
Domati[17]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	8 / 11	Limited
Vironen[33]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	7 / 11	Limited
Sapp[37]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	7 / 11	Limited
Trentham-Dietz[38]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8 / 11	Limited
Wilson[34]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	9 / 11	Probably robust
Serpentini[27]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	9 / 11	Limited

First Author	Conceptual		Measurement			Methodology				Interpretation		Check list Score	Expected methodological quality
	A priori hypothesis stated (Yes/No/NA)	Rationale for instrument reported (Yes/No)	Psychometric properties reported (Yes/No)	Cultural validity verified (Yes/No/NA)	Adequacy of domains covered (Yes/No)	Instrument administration reported (Yes/No)	Baseline compliance reported (Yes/No)	Timing of assessments documented (Yes/No)	Missing data documented (Yes/No)	Clinical significance addressed (Yes/No)	Presentation of results in general (Yes/No)		
Hornbrook[39]	No	No	Yes	Yes	Yes	Yes	Yes	No	No	Yes	Yes	7 / 11	Limited
Giesinger[18]	No	No	Yes	No	Yes	Yes	No	No	No	Yes	Yes	5 / 11	Limited
Trninic[32]	No	Yes	Yes	No	Yes	No	Yes	Yes	Yes	No	Yes	7 / 11	Limited
Caravati-Jouvenceaux[15]	No	No	Yes	Yes	No	Yes	Yes	Yes	No	Yes	Yes	7 / 11	Limited
Gall[40]	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	8 / 11	Limited
Reeve[6]	No	No	Yes	Yes	Yes	Yes	No	Yes	No	Yes	Yes	7 / 11	Limited
Den Oudsten[16]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	9 / 11	Probably robust
Thong 2013[28]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	8 / 11	Limited
Orsini[23]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	7 / 11	Limited
Traa[31]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	7 / 11	Limited
Wong[42]	No	No	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes	7 / 11	Limited
Pollack[36]	No	No	Yes	No	Yes	Yes	Yes	Yes	Yes	No	Yes	7 / 11	Limited

Running title: Quality of CRC HRQOL comparative studies

Table 5. Further quality assessments of included studies

	Comparison Group				
First Author	**Compare with their populations**	**Contemporary comparison (Yes, ≤5yrs; No, >5yrs)**	**Source of comparison group**	**Matched comparison group**	**Matching criteria**
Arndt 2004[13]	Yes	Yes	German general population	No	Nil
Arndt 2006[14]	Yes	Yes	German general population	No	Nil
Jansen[20]	Yes	No	German general population	No	Nil
Kopp[21]	Yes	Yes	German general population	No	Nil
Rauch[26]	No	No	Norwegian and German general populations	No	Nil
Neuman[22]	No	No	German general population	No	Nil
Pucciarelli 2008[25]	No	No	German general population	Yes	Age- and sex-matching
Zajac[35]	No	No	German general population	No	Nil
Pucciarelli 2010[24]	No	No	German general population	Yes	Age- and sex-matching
Guren[19]	Yes	Yes	Norwegian general population	No	Nil
Thong 2011a[30]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Thong 2011b[29]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Austin[41]	Yes	Yes	Australian general population	No	Nil
Domati[17]	Yes	No	Italian general population	No	Nil
Vironen[33]	Yes	No	Finnish general population	No	Nil
Sapp[37]	Yes	Yes	US woman general population	No	Nil
Trentham-Dietz[38]	Yes	Yes	US woman general population	No	Nil
Wilson[34]	Yes	No	UK general population	No	Nil
Serpentini[27]	Yes	No	Italian general population	No	Nil

	Comparison Group				
First Author	**Compare with their populations**	**Contemporary comparison (Yes, ≤5yrs; No, >5yrs)**	**Source of comparison group**	**Matched comparison group**	**Matching criteria**
Hornbrook[39]	Yes	No	US general population	No	Nil
Giesinger[18]	Yes	Yes	Austrian general population	Yes	Age- and sex-matching
Trninic[32]	Yes	Yes	Healthy population	No	Nil
Caravati-Jouvenceaux[15]	Yes	Yes	French general population	Yes	Age-, sex- and residence area-matching
Gall[40]	Yes	Yes	Australian general population	No	Nil
Reeve[6]	Yes	Yes	Individuals without cancer	Yes	Propensity score matching
Den Oudsten[16]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Thong 2013[28]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Orsini[23]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Traa[31]	Yes	Yes	Dutch general population	Yes	Age- and sex-matching
Wong[42]	Yes	Yes	Hong Kong general population	Yes	Age- and sex-matching
Pollack[36]	Yes	Yes	Sweden general population	Yes	Age-matching

Figure 1. PRISMA Flow Diagram of the literature search and selection process