

How health-related quality of life assessment should be used in advanced colorectal cancer clinical trials

F. Bonnetain,¹ C. Borg,² R. R. Adams,³ J. A. Ajani,⁴ A. Benson,⁵ H. Bleiberg,⁶ B. Chibaudel,⁷ E. Diaz-Rubio,⁸ J. Y. Douillard,⁹ C. S. Fuchs,¹⁰ B. J. Giantonio,¹¹ R. Goldberg,¹² V. Heinemann,¹³ M. Koopman,¹⁴ R. Labianca,¹⁵ A. K. Larsen,¹⁶ T. Maughan,¹⁷ E. Mitchell,¹⁸ M. Peeters,¹⁹ C. J. A. Punt,²⁰ H. J. Schmoll,²¹ C. Tournigand,²² A. de Gramont,⁷

¹Methodology and Quality of Life Unit, Oncology department (INSERM UMR 1098); Quality of Life and Cancer Clinical Research Platform, University Hospital of Besançon, Besançon, France;

²Department of Medical Oncology, University Hospital of Besançon; Centre investigation Clinique en biothérapie, CIC-1431; 11UMR1098 INSERM/Université de Franche Comté/Etablissement Français du Sang; Department of Oncology, University Hospital of Besançon, Besançon, France;

³Cardiff University and Velindre Cancer Centre, Cardiff, UK;

⁴Department of Gastrointestinal Medical Oncology, The University of Texas MD Anderson Cancer Center, Houston, Texas, USA;

⁵Division of Hematology/Oncology, Northwestern Medical Group, Chicago, USA;

⁶Montagne de Saint Job, Brussels, Belgium;

⁷Institut Hospitalier Franco-Britannique, Levallois-Perret, France;

⁸Medical Oncology Department, Hospital Clínico San Carlos, Madrid, Spain;

⁹Medical Oncology, Institut de Cancérologie de l'Ouest (ICO), Nantes St-Herblain, France;

¹⁰Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, USA;

¹¹Division of Hematology/Oncology, Department of Medicine, University of Pennsylvania, Philadelphia, USA;

¹²Department of Medicine, The Ohio State University Comprehensive Cancer Center-James Cancer Hospital and Solove Research Institute, Columbus, Ohio, USA;

¹³Department of Internal Medicine III and Comprehensive Cancer Center, Klinikum Grosshadern, Ludwig-Maximilians-University of Munich, Munich, Germany;

¹⁴Department of Medical Oncology, University Medical Center Utrecht, Utrecht, The Netherlands;

¹⁵Cancer Center, Ospedale Giovanni XXIII, Bergamo, Italy;

¹⁶Cancer Biology and Therapeutics, INSERM and Université Pierre et Marie Curie, Kourilsky building, Hôpital Saint-Antoine, Paris, France;

¹⁷CRUK/MRC Oxford Institute for Radiation Oncology, Gray Laboratories, University of Oxford, Oxford, UK;

¹⁸Department of Medical Oncology, Thomas Jefferson University, Philadelphia, USA;

¹⁹Center for Oncological Research Antwerp, University of Antwerp; Department of Oncology, Antwerp University Hospital, Antwerp, Belgium;

²⁰Department of Medical Oncology, Academic Medical Centre, University of Amsterdam, Amsterdam, The Netherlands;

²¹Department of Internal Medicine IV, University Clinic Halle, Martin-Luther-University Halle-Wittenberg, Halle, Germany;

²²Department of Oncology, University of Paris Est Creteil; APHP, Henri-Mondor Hospital, Créteil, France

Running title: Health-related quality of life in advanced colorectal cancer

Corresponding author:

Prof Franck Bonnetain,

Methodology and Quality of Life Unit, Oncology Department (INSERM UMR 1098);
Quality of Life and Cancer Clinical Research Platform, 3 Boulevard Fleming, 25030
Besançon, France; Tel: 00 33 (0) 3 81 21 92 06; Email: franck.bonnetain@univ-fcomte.fr/
fbonnetain@chu-besancon.fr

Word count: 3989

The total number of tables: 1 Table + 2 Supplementary tables online only

Key message

Health-related quality of life (HRQoL) is increasingly used as an endpoint in cancer clinical trials of the new therapeutic strategies. It provides added and more accurate predictive and prognostic value to other clinical parameters. HRQoL should be also considered as primary/co-primary endpoint in the advanced disease setting to improve the decision-making process at the individual patient level.

Abstract

Traditionally, the efficacy of cancer treatment in patients with advance or metastatic disease in clinical studies has been studied using overall survival (OS) and more recently tumor-based endpoints such as progression-free survival, measurements of response to treatment. However, these seem not to be the relevant clinical endpoints in current situation if such endpoints were no validated as surrogate of OS to demonstrate the clinical efficacy. Appropriate, meaningful, primary patient-oriented and patient-reported endpoints that adequately measure the effects of new therapeutic interventions are then crucial for the advancement of clinical research in metastatic colorectal cancer to complement the results of tumor-based endpoints. Health-related quality of life (HRQoL), is effectively an evaluation of quality of life and its relationship with health over time. HRQoL includes the patient report at least of the way a disease or its treatment affects its physical, emotional and social well-being. Over the past few years, several phase III trials in a variety of solid cancers have assessed the incremental value of HRQoL in addition to the traditional endpoints of tumor response and survival results. HRQoL could provide not only complementary clinical data to the primary outcomes, but also more precise predictive and prognostic value. This endpoint is useful for both clinicians and patients in order to achieve the dogma of precision medicine.

The present article examines the use of HRQoL in phase III metastatic colorectal cancer clinical trials, outlines the importance of HRQoL assessment methods, analysis, and results presentation. Moreover, it discusses the relevance of including HRQoL as a primary/co-primary endpoint to support the progression-free survival results and to assess efficacy of treatment in the advanced disease setting.

Key words: Health-related quality of life, longitudinal, methodology, colorectal cancer, clinical trial, endpoint

Introduction

The choice of a primary endpoint reflecting the true clinical benefit to the patient is paramount in clinical phase III trial design. Although, overall survival (OS) has been considered as the most relevant endpoint, it is now widely recognized as challenging when several therapeutic lines are available.[1] Other tumor-related outcomes such as progression-free survival (PFS) and response rate (RR) in the metastatic setting and disease-free survival (DFS) or recurrence-free survival in the adjuvant setting have been used.[2] These tumor-related endpoints, however, are not necessarily validated surrogates for OS nor do they translate into significant improvements in the duration and/or quality of survival. In this context, the value of several new therapeutic treatments has been called into question. Some interventions that provide only a marginal benefit as judged by extension of PFS, for example, might not be of sufficient value to patients, specifically those with advanced-stage cancer.

Over the past four decades, health-related quality of life (HRQoL) has been integrated into phase III clinical trials, including metastatic colorectal cancer (mCRC) studies, as important clinically relevant patient-reported outcome (PRO). PROs, including HRQoL, symptoms, treatment preferences, and patient satisfaction reflect patient subjective disease evaluation and treatment effect on patient daily life. HRQoL has been found a key indicator of patient treatment efficacy and safety by meta-analysis.[3-7] Better QoL physical function and global QoL score were independent predictors of longer OS, while better cognitive QoL was predictive of less adverse events in patients with pancreatic cancer.[8]

Despite the above mentioned findings and the fact that the US Food and Drug Administration (FDA) recognized HRQoL as a component endpoint for cancer therapy approvals [9], HRQoL is frequently neither adequately captured in clinical trials of advanced CRC (inaccurately applied, used as secondary/tertiary endpoint, estimated in post-hoc analysis) nor translated (or rarely) into clinical decision-making.

We outline a potential added value of HRQoL in the advanced CRC setting and discuss selected examples of phase III CRC trials that highlight the relevance of including HRQoL as a primary/co-primary measure of new therapeutic interventions. The methodological quality of HRQoL assessment, analysis, and presentation are also discussed.

Status of HRQoL

In mCRC clinical phase III trials, OS has been considered the optimal measure of treatment benefit due to its objectivity and accuracy. However it requires a long-term follow-up and a large number of patients to show statistically significant differences between treatments. Therefore, PFS is used as surrogate for OS, particularly in studies evaluating novel cancer therapies targeting a molecularly defined specific subpopulation of patients. Unlike OS, PFS is subjected to inherent biases and to measurement error. In addition, PFS may not reflect a true clinical benefit and it may not always be clinically relevant, especially when marginal treatment effects do not translate into OS increase. Large differences in PFS between the trials arms may then be required in order to consider it as relevant indicator of efficacy. Surrogacy of PFS for OS with fluoropyrimidines alone is weaker with currently available lines of treatment based on new drugs and targeted therapies.[1] Interestingly, the results of a survey conducted among clinicians and methodologists involved in the conduct of digestive cancer trials showed that in mCRC, PFS (75%), QoL together with PFS and R0 metastatic resection rate (56% both) were considered the ‘optimal’ alternative endpoints of OS in the metastatic setting.[10] Association between PFS and HRQoL was found in CRC patients treated with panitumumab.[11] Siena et al.,[11] reported that a lack of disease progression at week 8 was associated with significantly and clinically meaningful lower CRC symptomatology (National Comprehensive Cancer Network [NCCN]/ Functional Assessment of Cancer Therapy [FACT] CRC symptom index [FCSI] and HRQoL (EuroQol-5D [EQ-5D]))

for panitumumab and best supportive care patients and higher HRQoL for panitumumab patients only. However, this study was performed in pre-treated patients who had failed on all available therapies at that time, a setting where symptoms are frequent and HRQoL is often already impaired at the time of study enrollment, which is usually not the case in mCRC when metastasis are first diagnosed. In earlier disease stages, showing that a new drug that improves PFS did not compromise HRQoL, reinforce the survival results and influence the assessment value of the intervention.[12]

Given the above findings, any increase in survival, specifically in PFS only without an OS benefit, should be accompanied by a gain and/or an improved or maintained HRQoL if it is to be considered clinically relevant.[13, 14] Therefore, we propose the HRQoL assessment alongside with PFS and other potential endpoints in order to acquire a clinically reliable net value. HRQoL could also have a greater significance with intermittent therapy, in which PFS in its purest form is probably unhelpful and HRQoL is likely improved. In the palliative setting for mCRC, a goal of strategies with recent FDA approved agents that show small relative OS benefit, would be to maintain HRQoL instead of prolonging OS.

Assessment of HRQoL

HRQoL is commonly assessed using patient-reported questionnaires. A wide range of generic and cancer-specific modules to its measure have been developed and refined in the last years. The EORTC QOL Core Questionnaire (EORTC QOL-C30),[15] the Functional Assessment of Cancer Therapy-General (FACT-G),[16] the 36-item short-form health survey (SF-36)[17], and the EuroQoL-5D[18] are the most widely used.

Colorectal cancer and its treatments may adversely affect different dimensions of HRQoL, especially physical (e.g. social limitations as a result of physical health, pain/discomfort, general health perception), social (e.g. distress management, inability to socialize), economic

(e.g. transportation/medical costs, missed employment), emotional (e.g. anxiety, depression, low self-esteem), family (e.g. inability to provide appropriate support) and medical aspects (e.g. diarrhea, fatigue, impaired body image, sexual problems). These factors may affect treatment decisions, keeping medical appointments, following the recommended regimen/recommended medications/nutrition, or reporting adverse events or other factors to providers. Therefore, the choice of a tool to assess HRQoL is important in trials for advanced cancer to capture potential and relevant facts about the short and long-term impacts/effects of treatment and the cancer.

The FACT-C (discriminates colon vs. rectal patients) consisting of the 28-item FACT-G and the 9-item colon subscale is the most commonly used CRC specific questionnaire, validated in many countries.[19-23]. Two specific modules, the EORTC QLQ-CR29[24] and the liver metastasis EORTC QLQ-LMC21[25], were designed for use with the EORTC QLQ-C30 in CRC. Few data are available in the literature on the responsiveness of these questionnaires to change (e.g. the spread of cancer, chemotherapy outcomes).[22, 26] The EORTC QLQ-CR29 is psychometrically validated in colon and rectal cancer patients at various stages of their treatment that covers physical, psychosocial,, and CRC-specific symptom-oriented questions. This questionnaire has sufficient validity and reliability to supplement the EORTC QLQ-C30.[24] The validated EORTC QLQ-LMC21, covers symptom and function items specific to patients with CRC liver metastases and measures activity problems, abdominal pain, nutritional problems, and emotional function.[25] The low anterior resection score (LARS) is a 5-item questionnaire for evaluation of HRQoL in patients following surgery for rectal cancer. Validation of the LARS score is in progress.[27]

Some instruments such as the National Institute of Health's (NIH) Patient-Reported Outcome Measurement Information System (PROMIS) were developed to be applicable across diseases and not to be used as cancer-specific tool.[28, 29] PROMIS, with standardized PRO measures

referenced to the US general population, covers both generic and more specific items (e.g. sexual function). It integrates the use of item response theory and computer adaptive testing to create individualized questionnaires.[28, 29] The PROMIS relative performance psychometrics and, its prognostic nature for OS, DFS, and time to progression as well as the PRO-Common Toxicity Criteria measures versus simple a single-item numerical analogue scale assessment in the context of clinical studies have been investigated in the US to facilitate adoption by clinical researchers.

HRQoL analysis and interpretation

Efficient, reliable, and clinically meaningful HRQoL assessment in the clinical and routine clinical practice setting is not without specific limitations and challenges. Several difficulties such as the utility and availability of HRQoL instruments, methodological concerns (e.g. validity/reliability of measures) and logistic/practical considerations (e.g. feasibility of tools, scoring, interpretation) are encountered. HRQoL questionnaires are often lengthy, not easily understandable, and with complex scoring algorithm making their administration for routine use overdue and difficult. Implementation of in-clinic information technology (e.g. computer tablets) for data collection could facilitate this process. Basch and colleagues showed that systematic web-based reporting of patient-reported symptoms with automated clinician e-mail alerts resulted in better HRQoL, fewer emergency room visits, fewer hospitalizations, a longer duration of palliative chemotherapy, OS, and quality-adjusted survival in patients treated with chemotherapy for metastatic breast, genitourinary, gynecologic, or lung cancers.[30]

Given the longitudinal nature of HRQoL studies, identification of the appropriate time-points for clinically meaningful changes is also challenging. In addition, differences in the HRQoL definition perception, non-standardized and inconsistent use of multiple instruments, and

differences in reporting make a cross-study comparison difficult. Moreover, the HRQoL results interpretation can be problematic due to missing data and the patient dropout (health status deterioration, progression/relapse, death).[31] Patient response-shifts over time further complicate the HRQoL data interpretation.[32]

Another challenging aspect is HRQoL longitudinal data analysis that remains largely unstandardized often due to a lack of uniform approaches. Moreover, this analysis frequently produces clinically insignificant results. Although several methods to analyze longitudinal HRQoL data are available, these are rarely used by the health professionals mainly due to their complexity. It was well highlighted by the AVAglio[33] and RTOG 0825[34] trials evaluating the addition of bevacizumab to radiotherapy in glioblastoma patients. While PFS was significantly different in favor of bevacizumab, no OS benefits were observed in either trial. However, HRQoL results were conflicting. In RTOG 0825, HRQoL was lower among bevacizumab-treated patients, while in AVAglio, the time to HRQoL score deterioration was longer in the bevacizumab group.

Hamidou et al., used time until definitive deterioration (TUDD) statistical approach[35, 36] in the analysis of patients with mCRC and showed that time to deterioration in HRQoL did not differ significantly according to type of treatment. TUDD produces meaningful longitudinal HRQoL results easily interpreted by clinicians and thus improving clinical decision-making.

HRQoL relationship to OS

Prognostic value

HRQoL, besides being an important endpoint in clinical trials, is also an independent predictor of OS and, to some extent, of the patient response to treatment (marker of disease progression).[4, 37, 38] Gotay et al.,[4] reported that in 36 of the 39 analyzed studies PROs were significant predictors of survival, often a more accurate predictor than ECOG performance status. Several studies have demonstrated that baseline HRQoL is associated with OS duration in CRC. In a pooled analysis from four randomized trials by Maisey et al., [37], better baseline global HRQoL scores (up to few months after treatment) were independently associated with longer survival in mCRC patients (> 1 year). The authors recommended routine HRQoL measurements for stratification and comparison of outcomes in different cohorts of patients. In two other studies by Braun et al., [38] and Efficace et al.,[39], a 10-point increase on the social functioning scale or in baseline global HRQoL score was associated with a 7% and 6% decrease in the patient's hazard of death, respectively. Some authors showed that at least one HRQoL domain supplements prognostic information in addition to clinical and socio-demographic variables (WHO performance status, distant metastases, age, sex, and clinical factors).[40, 41] Baseline physical functioning score was independently associated with OS in study of mCRC patients treated with oxaliplatin-based first-line chemotherapy,[42] while self-reported HRQoL was not a significant prognostic factor for CRC recurrence in other study.[43] In addition, in a trial of CRC patients with liver metastases, HRQoL offered more accurate prognostic information for OS than the number/volume of metastases.[44]

Effect of treatment on HRQoL

Introduction of biologic agents into mCRC first-line treatment has resulted in increased RRs, PFS, and OS.[45-51] Many of these therapies are combined with cytotoxic regimens (e.g. capecitabine, irinotecan, oxaliplatin) and thus may have a significant impact on the patients' HRQoL.

In patients with surgically unresectable metastatic disease, chemotherapy, given its modest influence on OS, is usually administered with palliative intent. Although the goal of palliative therapy is the quality of survival, the impact of anticancer treatment has been evaluated by the length of survival and toxicity. The main goal of chemotherapy, however, is to delay HRQoL deterioration by reduction of the severity of disease-related symptoms.

A treatment can be chosen if it significantly improves HRQoL and well-being of the patient without significant OS benefit versus comparable treatment. In the phase III COIN trial[51] intermittent chemotherapy was non-inferior to continuous treatment in terms of OS. However, subgroup analysis demonstrated that patients with normal baseline platelet counts could gain the benefits of intermittent chemotherapy without detriment in OS, are likely to experience fewer side effects, and have improved HRQoL; those with increased platelet levels seem to do significantly worse on intermittent treatment and should not receive treatment breaks.

A treatment may not be optional if symptoms and HRQoL remain similar to that under the comparable treatment, without significant OS improvement. The NORDIC-VII phase III study of mCRC treated with cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) or FLOX alone in did not show any OS, PFS, and QoL benefit with the addition of the biologic, even in patients who were deemed KRAS wild-type.[52]

Clear treatment decisions seem challenging when treatment offers better chances of survival, but with impaired HRQoL compared to conservative therapy or lower changes of OS, but improved HRQoL. In such cases, if the primary endpoints are not met, HRQoL outcomes are

critical to decision whether combination therapy is beneficial from the patients perspective. These examples show how HRQoL could potentially guide therapy in the future if appropriately integrated into studies.

Also it might be the right time to consider HRQoL for prognostic scores to optimize any therapeutic strategy efficacy and to consider the effects of a treatment strategy on HRQoL endpoint in order to be in line with a dogma of precision/personalized medicine.

Minimal clinically important difference (MCID)

With the expanded use of HRQoL endpoint and the increasing number of HRQoL instruments, it is necessary to interpret HRQoL in the context of MCID improvements. MCID defines the smallest change in HRQoL, going beyond statistical significance, beneficial from patients' perspective, and leading to a change of the patient care.[53] Anchor-based (HRQoL scores are correlated with another independent measure) and distribution-based (measure of variability/distribution of results) methods are used to capture MCIDs.

There is a body of the literature regarding the clinical significance of changes in different questionnaires. In the EORTC QLQ-C30, a 5-10 mean score difference is regarded as a small, but subjectively significant or clinically meaningful, a 10-20 point change is regarded as moderate, and a >20-point change as large.[54-56] Regarding CRC-specific questionnaires, Yost et al.,[57] identified MCID scores as: 1-2 points for the Colorectal Cancer Subscale (CCS) specific, 4-6 points for the Treatment Outcome Index colorectal, and 5-8 points for the FACT-C total. For the FCSI (sensitive to symptomatic progression in metastatic patients), the MCID was defined as a change in the range of 1.5-3.0 points.[58]

Response Evaluation Criteria In Solid Tumors (RECIST) criteria and HRQoL

RECIST criteria for HRQoL can be helpful to characterize deterioration/improvement of and the magnitude of change in HRQoL parameters that have clinically meaningful effects. Moreover, these can help to reach standardized longitudinal HRQoL analysis using the TUDD method. With the objective of OS and HRQoL improvement, as for tumor response and progression, RECIST criteria for HRQoL should be defined and integrated into trial design. Development and validation of RECIST criteria for HRQoL in mCRC will be addressed by the international ARCAD (Aide et Recherche en Cancerologie Digestive) group in the near future.[59-61] An ongoing project granted by the EORTC QOL group would be also done for all cancer localizations.

HRQoL in mCRC phase III clinical trials

We have identified 31 pivotal first and second-line randomized phase III clinical trials on anticancer therapies for mCRC with HRQoL endpoint (Supplementary Table S1, available at *Annals of Oncology* online). To pinpoint those trials we conducted a literature review including a PubMed search using the search terms “QoL”, “endpoint”, “phase III”, “colorectal”, “first/second-line “ and “randomized”; and an abstracts search from conference (ASCO, ASCO GI, ESMO) proceedings. Studies included mostly first-line therapy trials ($n = 29$; 93.5%). The majority of trials used HRQoL as pre-specified secondary endpoint (93.5%), except for two studies, in which HRQoL was adapted as a co-primary (alongside PFS) or tertiary objective. These observation points out that although the importance of HRQoL assessment in clinical trials is now generally accepted, many clinicians are still skeptical about its true additive value when testing different therapeutic strategies and are hesitant to accept HRQoL as a primary/co-primary endpoint. Furthermore, HRQoL data were presented separately for eight trials (25.8%). Although splitting up HRQoL data from survival results seems to be opportunity for a comprehensive way of reporting, such action has also certain

disadvantages. Separate reporting of survival and HRQoL results may reduce their value in clinical decision-making as clinicians are unlikely to read or are not aware of the successive paper published following the main results.[62, 63] It is therefore recommended that clinical and HRQoL sets of data are published in the same manuscript in order to provide more complete account of outcomes and in turns optimal therapeutic strategies in patients with advanced disease. Another alternative could be to publish a companion paper in the same journal at the same time, as reflected by the published results from the AURELIA phase III study of bevacizumab in ovarian cancer.[64, 65]

Although different measures were used to assess HRQoL in the selected studies, consistent use of the generic questionnaires was observed: EORTC QLQ-C30 (22/31; 71.0%), EQ-5D (7/31; 22.6%), and SF-36 (2/31; 6.4%). However, only eight of 31 reviewed trials (25.8%) used CRC-specific questionnaires (of which four used FACT-C, two used FACT-FCSI, one used EORTC QLQ-CR29, two used CCS, and one used TOI-C; Supplementary Table S1, available at *Annals of Oncology* online). These were used either alone or in association with generic questionnaires. While the generic HRQoL questionnaires are well-validated and robust measures, they do not cover specific mCRC-related aspects and therefore may miss effects of new therapies on patients' HRQoL. It is hence vital to use CRC-sensitive/discriminate questionnaires (and score them along with generic instruments) in mCRC patients.

HRQoL data quality may be diminished by poor compliance in questionnaire completion as the study progress. Quinten et al. [40] showed that patients who did not complete the questionnaires may have poorer survival. A generation gap between patients (the average age of mCRC patients is 65 years) may also contribute to HRQoL data non-compliance given different approaches of capturing HRQoL (e.g. telephone, electronic, paper). Only 14 of 31 reviewed trials (45.2%) reported patient compliance for HRQoL questionnaires completion

(Supplementary Table S2, available at *Annals of Oncology* online). Four trials (28.6%) provided the total number of patients who completed the questionnaire at baseline, while three of these did not specify compliance rates during the trial. In all trials, the baseline compliance rates ranged from 84% to 98%. As expected, the post-baseline decline in completing questionnaires was observed. Moreover, differences in reporting missing data (not sufficient/no data at all) and longitudinal analysis were observed across the retrieved studies. Missing data at later time points may be related to treatment-related toxicities, tumor progression, difficulties in completing questionnaires, or low stress on HRQoL measure. This aspect can introduce bias selection (towards worse or better HRQoL), reduction in statistical power to detect MCID (reduced number of observations), and finally lead to misleading results regarding HRQoL, which is particularly relevant in patients with advanced and progressive disease.

We also investigated the relationships between survival, tolerance, and the HRQoL scores (Table 1) across the reviewed trials. A significant difference between arms with regard to HRQoL was found in 16 out of 31 trials (51.6%), eight and five of which had PFS and OS as the primary endpoint, respectively, one had OS and DFS as the co-primary endpoints, one had OS and PFS as the co-primary endpoints, and one had time to failure strategy as the primary endpoint. Six of 15 trials (40%) and eight of 14 trials (57.1%) that demonstrated or did not show OS or PFS benefit, respectively, and two trials (100%) with OS negative results found significant differences in QoL scores. These observations are of importance as demonstrate that 50% of trials that did not show survival differences demonstrated significant HRQoL changes. Moreover, in only one trial,[66] the improved survival came at a cost of decreased HRQoL. Although greater tolerance was seen, generally it did not have any adverse impact on neither survival nor HRQoL.

Longitudinal HRQoL studies can reveal insight into changes in HRQoL scores as patients undergo treatment. Despite many sophisticated statistical approaches available, longitudinal analysis of HRQoL remains generally unstandardized and inconsistent, which can compromise the results between trials. Our assessment revealed remarkable across-trials longitudinal analysis heterogeneity. Although longitudinal designs were widely adopted, advanced statistical methods were commonly underutilized. The majority of studies were analyzed using traditional methods that account for repeated measures (20/30; 66.7%), such as pattern-mixture model, linear-mixed model, Wei-Lachin or Wilcoxon rank-sum tests. Only four of 30 studies (13.3%) implemented time-to-QoL deterioration (TTD), and only four used mixed-effect model for repeated measures approach. In six studies (20%) the statistical strategies were not at all reported (Supplementary Table S1, available at *Annals of Oncology* online). In addition, missing data was rarely discussed.

Suggestions for future mCRC trials

Considering the above, we propose the following for future mCRC clinical trials:

1. HRQoL should be strongly considered as a co-primary endpoint alongside a tumor-based parameters such as PFS in the first-line and later settings.[60, 67] Appropriate investment and strategies to ensure completion of data collection and minimization of drop out is critical.
2. HRQoL could be used as a composite endpoint with tumor parameters when: treatment effects on each component of endpoint is expected to be of similar efficacy, treatment is expected to have long-term negative effects on patients HRQoL, and treatment effects are estimated to be small, but clinically meaningful. This approach would ensure adequate sample size for meaningful HRQoL assessments.

3. Specific instruments such as the EORTC QLQ-C30 form associated with the EORTC QLQ-CR29 form should be used to increase the sensitivity for HRQoL changes detection and to compare between studies along with more general tools.

4. The specific HRQOL dimensions should be targeted to prevent multiple testing and inflation of type one error with associated decisions rules regarding the conclusion of the trial. Mandatory dimensions should include: global health status (true overall impact assessment of disease and its treatment on the patient, assessment of at least one dimension comparable across trials, impact measurement of other dimensions on global health), functional scales, symptoms scales (measurement of ongoing treatment and/or on treatment with late toxicity effects), and CRC-specific dimensions of the QLQ-CR 29 (capturing sensibility to change). Complementary dimensions could be targeted for exploratory purposes only. With HRQoL as a co-primary endpoint, complementary dimensions could be considered as secondary endpoints for confirmatory purposes.

5. Timing driven design should be preferred for collection of HRQoL data. Assessment times of HRQoL should coincide with the clinical care schedule dictated by the trial regimens, but it is essential to measure it at baseline (before randomization), at least every month until death for trials with an expected median OS ≤ 1 year, at least every month during treatment, and at least every 2-3 months until death for trials with an expected median OS > 1 year. For strategy trials, follow-up should be measured until death or at least treatment failure of the last chemotherapy line evaluated.

6. HRQoL should be an integrated component of prognostic scores (e.g. the GERCOR prognostic score)[68] and considered as a stratification factor in mCRC clinical trials.

Conclusion

Based on an extensive literature review and the experts opinions, we propose guidelines for the development of parameters that will facilitate the conduct of future mCRC clinical trials with HRQoL as a co-primary/composite endpoint and will ensure their methodological and analytical quality and comparability.

Acknowledgements

Magdalena Benetkiewicz (PhD) assisted the authors with editing the article; her work was funded by the Fondation A.R.CA.D.

This article was done within the framework of the Fondation A.R.CA.D activities; the authors received no financial support from the Fondation A.R.CA.D.

Acknowledged for their contribution and support:

USA: H. Hurwitz (Duke University Cancer Institute, Durham), S. Kopetz (M.D. Anderson Cancer Center, Houston);

Italy: F. Loupakis (U.O. Oncologia Medica 2, Azienda Ospedaliero-Universitaria Pisana, Istituto Toscano Tumori, Pisa);

France: C. Louvet (Institut Mutualiste Montsouris, Paris);

UK: M.T. Seymour (Cancer Medicine & Consultant Medical Oncologist, University of Leeds, Leeds);

Greece: J. Souglakos (University Hospital Heraklion, Crete);

Spain: J. Tabernero (Vall d'Hebron Institute of Oncology, Barcelona).

Funding

None declared. No grant number is applicable.

Disclosure

The authors have declared no conflicts of interest.

References

1. Shi Q, de Gramont A, Grothey A et al. Individual patient data analysis of progression-free survival versus overall survival as a first-line end point for metastatic colorectal cancer in modern randomized trials: findings from the analysis and research in cancers of the digestive system database. *J Clin Oncol* 2015; 33: 22-28.
2. Fiteni F, Westeel V, Pivot X et al. Endpoints in cancer clinical trials. *J Visc Surg* 2014; 151: 17-22.
3. Montazeri A. Quality of life data as prognostic indicators of survival in cancer patients: an overview of the literature from 1982 to 2008. *Health Qual Life Outcomes* 2009; 7: 102.
4. Gotay CC, Kawamoto CT, Bottomley A, Efficace F. The prognostic significance of patient-reported outcomes in cancer clinical trials. *J Clin Oncol* 2008; 26: 1355-1363.
5. Quinten C, Martinelli F, Coens C et al. A global analysis of multitrial data investigating quality of life and symptoms as prognostic factors for survival in different tumor sites. *Cancer* 2014; 120 302-311.
6. Tan AD, Novotny PJ, Kaur JS et al. A patient-level meta-analytic investigation of the prognostic significance of baseline quality of life (QOL) for overall survival (OS) among 3,704 patients participating in 24 North Central Cancer Treatment Group (NCCTG) and Mayo Clinic Cancer Center (MC) oncology clinical trials. 2007; 26: (suppl 1; abstr 9515).
7. Van Cutsem E, Moiseyenko VM, Tjulandin S et al. Phase III study of docetaxel and cisplatin plus fluorouracil compared with cisplatin and fluorouracil as first-line therapy for advanced gastric cancer: A report of the V325 study group. *J Clin Oncol* 2006; 24: 4991-4997.

8. Vickers MM, Tu D, Lee C et al. Significance of baseline quality of life scores in predicting clinical outcomes in an international phase III trial of advanced pancreatic cancer: NCIC CTG PA.3. 2013; 31 (suppl; abstr 4053).
9. US Department of Health and Human Services UFaDA, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research: Guidance for industry: Clinical trial endpoints for the approval of cancer drugs and biologics. <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm071590.pdf>. Accessed 21 May, 2014.
10. Methy N, Bedenne L, Bonnetain F. Surrogate endpoints for overall survival in digestive oncology trials: which candidates? A questionnaires survey among clinicians and methodologists. BMC Cancer 2010; 10: 277.
11. Siena S, Peeters M, Van Cutsem E et al. Association of progression-free survival with patient-reported outcomes and survival: results from a randomised phase 3 trial of panitumumab. Br J Cancer 2007; 97: 1469-1474.
12. Bennett L, Zhao Z, Barber B et al. Health-related quality of life in patients with metastatic colorectal cancer treated with panitumumab in first- or second-line treatment. Br J Cancer 2011; 105: 1495-1502.
13. Hamidou Z, Dabakuyo TS, Bonnetain F. Impact of response shift on longitudinal quality-of-life assessment in cancer clinical trials. Expert Rev Pharmacoecon Outcomes Res 2011; 11: 549-559.
14. Dabakuyo TS, Guillemin F, Conroy T et al. Response shift effects on measuring post-operative quality of life among breast cancer patients: a multicenter cohort study. Qual Life Res 2013; 22: 1-11.

15. Aaronson NK, Ahmedzai S, Bergman B et al. The European Organization for Research and Treatment of Cancer QLQ-C30: a quality-of-life instrument for use in international clinical trials in oncology. *J Natl Cancer Inst* 1993; 85: 365-376.
16. Cella DF, Tulsky DS, Gray G et al. The Functional Assessment of Cancer Therapy scale: development and validation of the general measure. *J Clin Oncol* 1993; 11: 570-579.
17. Ware JE, Jr., Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; 30: 473-483.
18. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; 37: 53-72.
19. Ward WL, Hahn EA, Mo F et al. Reliability and validity of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C) quality of life instrument. *Qual Life Res* 1999; 8: 181-195.
20. Yoo HJ, Kim JC, Eremenco S, Han OS. Quality of life in colorectal cancer patients with colectomy and the validation of the Functional Assessment of Cancer Therapy-Colorectal (FACT-C), Version 4. *J Pain Symptom Manage* 2005; 30: 24-32.
21. Rotonda C, Conroy T, Mercier M et al. Validation of the French version of the colorectal-specific quality-of-life questionnaires EORTC QLQ-CR38 and FACT-C. *Qual Life Res* 2008; 17: 437-445.
22. Conroy T, Mercier M, Bonnetterre J et al. French version of FACT-G: validation and comparison with other cancer-specific instruments. *Eur J Cancer* 2004; 40: 2243-2252.
23. Dapuerto JJ, Francolino C, Servente L et al. Evaluation of the Functional Assessment of Cancer Therapy-General (FACT-G) Spanish Version 4 in South America: classic psychometric and item response theory analyses. *Health Qual Life Outcomes* 2003; 1: 32.
24. 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. *Eur J Cancer* 2009; 45: 3017-3026.

25. Blazeby JM, Fayers P, Conroy T et al. Validation of the European Organization for Research and Treatment of Cancer QLQ-LMC21 questionnaire for assessment of patient-reported outcomes during treatment of colorectal liver metastases. *Br J Surg* 2009; 96: 291-298.
26. Osoba D, Zee B, Pater J et al. Psychometric properties and responsiveness of the EORTC quality of Life Questionnaire (QLQ-C30) in patients with breast, ovarian and lung cancer. *Qual Life Res* 1994; 3: 353-364.
27. Juul T, Ahlberg M, Biondo S et al. International validation of the low anterior resection syndrome score. *Ann Surg* 2014; 259: 728-734.
28. Cella D, Riley W, Stone A et al. The Patient-Reported Outcomes Measurement Information System (PROMIS) developed and tested its first wave of adult self-reported health outcome item banks: 2005-2008. *J Clin Epidemiol* 2010; 63: 1179-1194.
29. Riley WT, Rothrock N, Bruce B et al. Patient-reported outcomes measurement information system (PROMIS) domain names and definitions revisions: further evaluation of content validity in IRT-derived item banks. *Qual Life Res* 2010; 19: 1311-1321.
30. Basch E, Deal AM, Kris MG et al. Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial. *J Clin Oncol* 2016; 34: 557-565.
31. Post WJ, Buijs C, Stolk RP et al. The analysis of longitudinal quality of life measures with informative drop-out: a pattern mixture approach. *Qual Life Res* 2010; 19: 137-148.
32. Efficace F, Bottomley A. Do quality-of-life randomized clinical trials support clinicians in their decision-making? *J Clin Oncol* 2002; 20: 4126-4127.
33. Chinot OL, Wick W, Mason W et al. Bevacizumab plus radiotherapy-temozolomide for newly-diagnosed glioblastoma. *N Eng J Med* 2014; 370: 709-722.

34. Gilbert MR, Dignam J, Won M et al. RTOG 0825: Phase III double-blind placebo-controlled trial evaluating bevacizumab (Bev) in patients (Pts) with newly diagnosed glioblastoma (GBM). *J Clin Oncol* 2013; 31: (suppl; abstr 1).
35. Hamidou Z, Chibaudel B, Hebbar M et al. Time to Definitive Health-Related Quality of Life Score Deterioration in Patients with Resectable Metastatic Colorectal Cancer Treated with FOLFOX4 versus Sequential Dose-Dense FOLFOX7 followed by FOLFIRI: The MIROX Randomized Phase III Trial. *PLoS One* 2016; 11: e0157067.
36. Bonnetain F, Dahan L, Maillard E et al. Time until definitive quality of life score deterioration as a means of longitudinal analysis for treatment trials in patients with metastatic pancreatic adenocarcinoma. *Eur J Cancer* 2010; 46: 2753-2762.
37. Maisey NR, Norman A, Watson M et al. Baseline quality of life predicts survival in patients with advanced colorectal cancer. *Eur J Cancer* 2002; 38: 1351-1357.
38. Braun DP, Gupta D, Grutsch JF, Staren ED. Can changes in health related quality of life scores predict survival in stages III and IV colorectal cancer? *Health Qual Life Outcomes* 2011; 9: 62.
39. Efficace F, Innominato PF, Bjarnason G et al. Validation of patient's self-reported social functioning as an independent prognostic factor for survival in metastatic colorectal cancer patients: results of an international study by the Chronotherapy Group of the European Organisation for Research and Treatment of Cancer. *J Clin Oncol* 2008; 26: 2020-2026.
40. Quidde J, Arnold D, Hegewisch-Becker S et al. Quality of Life (QoL) in patients with metastatic colorectal cancer (mCRC) receiving maintenance therapy after first-line inductive treatment: A QoL sub-analysis of the AIO KKK 0207 phase III trial. *Annals of Oncology* 2014; 25 (suppl_4; iv167-iv209).

41. Mol L, Ottevanger PB, Koopman M, Punt CJ. The prognostic value of WHO performance status in relation to quality of life in advanced colorectal cancer patients. *Eur J Cancer* 2016; 66: 138-143.
42. Comella P, Casaretti R, Manzo R et al. Baseline physical functioning status of metastatic colorectal cancer patients predicts the overall survival but not the activity of a front-line oxaliplatin-fluoropyrimidine doublet. *Acta Oncol* 2010; 49: 50-56.
43. Wong CK, Law WL, Wan YF et al. Health-related quality of life and risk of colorectal cancer recurrence and All-cause death among advanced stages of colorectal cancer 1-year after diagnosis. *BMC Cancer* 2014; 14: 337.
44. Earlam S, Glover C, Fordy C et al. Relation between tumor size, quality of life, and survival in patients with colorectal liver metastases. *J Clin Oncol* 1996; 14: 171-175.
45. de Gramont A, Figer A, Seymour M et al. Leucovorin and fluorouracil with or without oxaliplatin as first-line treatment in advanced colorectal cancer. *J Clin Oncol* 2000; 18: 2938-2947.
46. Tournigand C, Andre T, Achille E et al. FOLFIRI followed by FOLFOX6 or the reverse sequence in advanced colorectal cancer: a randomized GERCOR study. *J Clin Oncol* 2004; 22: 229-237.
47. Heinemann V, von Weikersthal LF, Decker T et al. FOLFIRI plus cetuximab versus FOLFIRI plus bevacizumab as first-line treatment for patients with metastatic colorectal cancer (FIRE-3): a randomised, open-label, phase 3 trial. *Lancet Oncol* 2014; 15: 1065-1075.
48. Hurwitz H, Fehrenbacher L, Novotny W et al. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *N Engl J Med* 2004; 350: 2335-2342.
49. Douillard JY, Siena S, Cassidy J et al. Randomized, phase III trial of panitumumab with infusional fluorouracil, leucovorin, and oxaliplatin (FOLFOX4) versus FOLFOX4 alone

as first-line treatment in patients with previously untreated metastatic colorectal cancer: the PRIME study. *J Clin Oncol* 2010; 28: 4697-4705.

50. Douillard JY, Siena S, Cassidy J et al. Final results from PRIME: randomized phase III study of panitumumab with FOLFOX4 for first-line treatment of metastatic colorectal cancer. *Ann Oncol* 2014; 25: 1346-1355.

51. Adams RA, Meade AM, Seymour MT et al. Intermittent versus continuous oxaliplatin and fluoropyrimidine combination chemotherapy for first-line treatment of advanced colorectal cancer: results of the randomised phase 3 MRC COIN trial. *Lancet Oncol* 2011; 12: 642-653.

52. Tveit KM, Guren T, Glimelius B et al. Phase III trial of cetuximab with continuous or intermittent fluorouracil, leucovorin, and oxaliplatin (Nordic FLOX) versus FLOX alone in first-line treatment of metastatic colorectal cancer: the NORDIC-VII study. *J Clin Oncol* 2012; 30: 1755-1762.

53. Jaeschke R, Singer J, Guyatt GH. Measurement of health status. Ascertaining the minimal clinically important difference. *Control Clin Trials* 1989; 10: 407-415.

54. Osoba D, Rodrigues G, Myles J et al. Interpreting the significance of changes in health-related quality-of-life scores. *J Clin Oncol* 1998; 16: 139-144.

55. Ringash J, O'Sullivan B, Bezjak A, Redelmeier DA. Interpreting clinically significant changes in patient-reported outcomes. *Cancer* 2007; 110: 196-202.

56. Cocks K, King MT, Velikova G et al. Evidence-based guidelines for determination of sample size and interpretation of the European Organisation for the Research and Treatment of Cancer Quality of Life Questionnaire Core 30. *J Clin Oncol* 2011; 29: 89-96.

57. 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.

58. Colwell HH, Mathias SD, Turner MP et al. Psychometric evaluation of the FACT Colorectal Cancer Symptom Index (FCSI-9): reliability, validity, responsiveness, and clinical meaningfulness. *Oncologist* 2010; 15: 308-316.

59. Anota A, Hamidou Z, Paget-Bailly S et al. Time to health-related quality of life score deterioration as a modality of longitudinal analysis for health-related quality of life studies in oncology: do we need RECIST for quality of life to achieve standardization? *Qual Life Res* 2013.

60. Bonnetain F, Fiteni F, Efficace F, Anota A. Statistical Challenges in the Analysis of Health-Related Quality of Life in Cancer Clinical Trials. *J Clin Oncol* 2016; 34: 1953-1956.

61. Bottomley A, Pe M, Sloan J et al. Analysing data from patient-reported outcome and quality of life endpoints for cancer clinical trials: a start in setting international standards. *Lancet Oncol* 2016; 17: e510-e514.

62. Patrick D. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *Value Health* 2013; 16: 455-456.

63. Calvert M, Blazeby J, Altman DG et al. Reporting of patient-reported outcomes in randomized trials: the CONSORT PRO extension. *JAMA* 2013; 309: 814-822.

64. Poveda AM, Selle F, Hilpert F et al. Bevacizumab Combined With Weekly Paclitaxel, Pegylated Liposomal Doxorubicin, or Topotecan in Platinum-Resistant Recurrent Ovarian Cancer: Analysis by Chemotherapy Cohort of the Randomized Phase III AURELIA Trial. *J Clin Oncol* 2015; 33: 3836-3838.

65. Tomao F, Tomao S, Benedetti Panici P. Combination of bevacizumab and chemotherapy for platinum-resistant recurrent ovarian cancer: some observations about the AURELIA trial. *J Clin Oncol* 2014; 32: 3580.

66. Ringash J, Au HJ, Siu LL et al. Quality of life in patients with K-RAS wild-type colorectal cancer: the CO.20 phase 3 randomized trial. *Cancer* 2014; 120: 181-189.
67. Wilson MK, Karakasis K, Oza AM. Outcomes and endpoints in trials of cancer treatment: the past, present, and future. *Lancet Oncol* 2015; 16: e32-42.
68. Chibaudel B, Bonnetain F, Tournigand C et al. Simplified prognostic model in patients with oxaliplatin-based or irinotecan-based first-line chemotherapy for metastatic colorectal cancer: a GERCOR study. *Oncologist* 2011; 16: 1228-1238.
69. 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.
70. Tebbutt NC, Wilson K, Gebiski VJ et al. Capecitabine, bevacizumab, and mitomycin in first-line treatment of metastatic colorectal cancer: results of the Australasian Gastrointestinal Trials Group Randomized Phase III MAX Study. *J Clin Oncol* 2010; 28: 3191-3198.
71. Stockler M, Zannino D, Wilson K, et al. . Patient-rated outcomes (PRO) in a randomized trial of first-line chemotherapy with capecitabine (C), bevacizumab (B), and mitomycin-C (M) for metastatic colorectal cancer: The AGITG MAX trial. *J Clin Oncol* 2010; 28 (suppl e14036).
72. Koopman M, Antonini NF, Douma J et al. Sequential versus combination chemotherapy with capecitabine, irinotecan, and oxaliplatin in advanced colorectal cancer (CAIRO): a phase III randomised controlled trial. *Lancet* 2007; 370: 135-142.
73. Koopman M SL, May A, et al. . Final results and subgroup analyses of the phase 3 CAIRO3 study: maintenance treatment with capecitabine and bevacizumab versus observation after induction treatment with chemotherapy and bevacizumab in metastatic colorectal cancer (mCRC). *J Clin Oncol* 2014;32 (suppl 3; abstr LBA388).

74. Punt CJA, Simkens LHJ, May A. Updated results including quality of life of the phase III CAIRO3 study of the Dutch Colorectal Cancer Group (DCCG): maintenance treatment with capecitabine and bevacizumab versus observation after induction treatment with chemotherapy and bevacizumab in metastatic colorectal cancer (mCRC). *Eur J Cancer* 2013; 49 (2 suppl; abstr S486).
75. Simkens LH, van Tinteren H, May A et al. Maintenance treatment with capecitabine and bevacizumab in metastatic colorectal cancer (CAIRO3): a phase 3 randomised controlled trial of the Dutch Colorectal Cancer Group. *Lancet* 2015; 385: 1843-1852.
76. Grothey A, Van Cutsem E, Sobrero A et al. Regorafenib monotherapy for previously treated metastatic colorectal cancer (CORRECT): an international, multicentre, randomised, placebo-controlled, phase 3 trial. *Lancet* 2013; 381: 303-312.
77. Sobrero AF, Maurel J, Fehrenbacher L et al. EPIC: phase III trial of cetuximab plus irinotecan after fluoropyrimidine and oxaliplatin failure in patients with metastatic colorectal cancer. *J Clin Oncol* 2008; 26: 2311-2319.
78. Naughton MJ, Schrag D, Venook AP et al. Quality of life (QOL) and toxicity among patients in CALGB 80405. *J Clin Oncol* 2013; 31 (suppl; abstr 3611).
79. Venook AP, Niedzwiecki D, Lenz HJ et al. Phase III trial of irinotecan/5-FU/leucovorin (FOLFIRI) or oxaliplatin/5-FU/leucovorin (mFOLFOX6) with bevacizumab or cetuximab for patients with KRAS wild-type untreated metastatic adenocarcinoma of the colon or rectum. *J Clin Oncol* 2014; 32 (suppl; abstr LBA3).
80. Tol J, Koopman M, Cats A et al. Chemotherapy, bevacizumab, and cetuximab in metastatic colorectal cancer. *N Engl J Med* 2009; 360: 563-572.
81. Lembersky BC, Wieand HS, Petrelli NJ et al. Oral uracil and tegafur plus leucovorin compared with intravenous fluorouracil and leucovorin in stage II and III carcinoma of the

colon: results from National Surgical Adjuvant Breast and Bowel Project Protocol C-06. *J Clin Oncol* 2006; 24: 2059-2064.

82. Kopec JA, Yothers G, Ganz PA et al. Quality of life in operable colon cancer patients receiving oral compared with intravenous chemotherapy: results from National Surgical Adjuvant Breast and Bowel Project Trial C-06. *J Clin Oncol* 2007; 25: 424-430.

83. Kemeny NE, Niedzwiecki D, Hollis DR et al. Hepatic arterial infusion versus systemic therapy for hepatic metastases from colorectal cancer: a randomized trial of efficacy, quality of life, and molecular markers (CALGB 9481). *J Clin Oncol* 2006; 24: 1395-1403.

84. Hegewisch-Becker S, Graeven U, Lerchenmuller CA et al. Maintenance strategies after first-line oxaliplatin plus fluoropyrimidine plus bevacizumab for patients with metastatic colorectal cancer (AIO 0207): a randomised, non-inferiority, open-label, phase 3 trial. *Lancet Oncol* 2015; 16: 1355-1369.

85. Arnold D, Graeven U, Lerchenmuller CA et al. Maintenance strategy with fluoropyrimidines (FP) plus Bevacizumab (Bev), Bev alone, or no treatment, following a standard combination of FP, oxaliplatin (Ox), and Bev as first-line treatment for patients with metastatic colorectal cancer (mCRC): A phase III non-inferiority trial (AIO KRK 0207). *J Clin Oncol* 32 (suppl; abstr 3503).

86. Maughan TS, James RD, Kerr DJ et al. Comparison of survival, palliation, and quality of life with three chemotherapy regimens in metastatic colorectal cancer: a multicentre randomised trial. *Lancet* 2002; 359: 1555-1563.

87. Van Cutsem E, Kohne CH, Hitre E et al. Cetuximab and chemotherapy as initial treatment for metastatic colorectal cancer. *N Engl J Med* 2009; 360: 1408-1417.

88. Van Cutsem E, Kohne CH, Lang I et al. Cetuximab plus irinotecan, fluorouracil, and leucovorin as first-line treatment for metastatic colorectal cancer: updated analysis of overall

survival according to tumor KRAS and BRAF mutation status. *J Clin Oncol* 2011; 29: 2011-2019.

89. Lang I, Kohne CH, Folprecht G et al. Quality of life analysis in patients with KRAS wild-type metastatic colorectal cancer treated first-line with cetuximab plus irinotecan, fluorouracil and leucovorin. *Eur J Cancer* 2013; 49: 439-448.

90. Falcone A, Ricci S, Brunetti I et al. Phase III trial of infusional fluorouracil, leucovorin, oxaliplatin, and irinotecan (FOLFOXIRI) compared with infusional fluorouracil, leucovorin, and irinotecan (FOLFIRI) as first-line treatment for metastatic colorectal cancer: the Gruppo Oncologico Nord Ovest. *J Clin Oncol* 2007; 25: 1670-1676.

91. 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. *J Clin Oncol* 2009; 27: 1822-1828.

92. Jonker DJ, O'Callaghan CJ, Karapetis CS et al. Cetuximab for the treatment of colorectal cancer. *N Engl J Med* 2007; 357: 2040-2048.

93. Siu LL, Shapiro JD, Jonker DJ et al. Phase III randomized, placebo-controlled study of cetuximab plus brivanib alaninate versus cetuximab plus placebo in patients with metastatic, chemotherapy-refractory, wild-type K-RAS colorectal carcinoma: the NCIC Clinical Trials Group and AGITG CO.20 Trial. *J Clin Oncol* 2013; 31: 2477-2484.

94. Peeters M, Price TJ, Cervantes A et al. Final results from a randomized phase 3 study of FOLFIRI {+/-} panitumumab for second-line treatment of metastatic colorectal cancer. *Ann Oncol* 2014; 25: 107-116.

95. Peeters M, Price TJ, Cervantes A et al. Randomized phase III study of panitumumab with fluorouracil, leucovorin, and irinotecan (FOLFIRI) compared with FOLFIRI alone as second-line treatment in patients with metastatic colorectal cancer. *J Clin Oncol* 2010; 28: 4706-4713.

96. Seymour MT, Brown SR, Middleton G et al. Panitumumab and irinotecan versus irinotecan alone for patients with KRAS wild-type, fluorouracil-resistant advanced colorectal cancer (PICCOLO): a prospectively stratified randomised trial. *Lancet Oncol* 2013; 14: 749-759.
97. Twelves C, Wong A, Nowacki MP et al. Capecitabine as adjuvant treatment for stage III colon cancer. *N Engl J Med* 2005; 352: 2696-2704.
98. Kohne CH, Wils J, Lorenz M et al. Randomized phase III study of high-dose fluorouracil given as a weekly 24-hour infusion with or without leucovorin versus bolus fluorouracil plus leucovorin in advanced colorectal cancer: European organization of Research and Treatment of Cancer Gastrointestinal Group Study 40952. *J Clin Oncol* 2003; 21: 3721-3728.
99. Saltz LB, Cox JV, Blanke C et al. Irinotecan plus fluorouracil and leucovorin for metastatic colorectal cancer. Irinotecan Study Group. *N Engl J Med* 2000; 343: 905-914.
100. Conroy T, Hebbar M, Bennouna J et al. Quality-of-life findings from a randomised phase-III study of XELOX vs FOLFOX-6 in metastatic colorectal cancer. *Br J Cancer* 2010; 102: 59-67.
101. Ducreux M BJ, Hebbar M, et al. . Efficacy and safety findings from a randomized phase III study of capecitabine (X) + oxiplatin (O) (XELOX) vs. infusional 5-FU/LV + O (FOLFOX-6) for metastatic colorectal cancer (MCRC). *J Clin Oncol* 2007;25 (suppl 4029).
102. Seymour MT, Thompson LC, Wasan HS et al. Chemotherapy options in elderly and frail patients with metastatic colorectal cancer (MRC FOCUS2): an open-label, randomised factorial trial. *Lancet* 2011; 377: 1749-1759.
103. Haller DG, Rothenberg ML, Wong AO et al. Oxaliplatin plus irinotecan compared with irinotecan alone as second-line treatment after single-agent fluoropyrimidine therapy for metastatic colorectal carcinoma. *J Clin Oncol* 2008; 26: 4544-4550.

104. Price TJ, Peeters M, Kim TW et al. Panitumumab versus cetuximab in patients with chemotherapy-refractory wild-type KRAS exon 2 metastatic colorectal cancer (ASPECCT): a randomised, multicentre, open-label, non-inferiority phase 3 study. *Lancet Oncol* 2014; 15: 569-579.
105. Van Cutsem E, Peeters M, Siena S et al. Open-label phase III trial of panitumumab plus best supportive care compared with best supportive care alone in patients with chemotherapy-refractory metastatic colorectal cancer. *J Clin Oncol* 2007; 25: 1658-1664.
106. Schmoll HJ, Cunningham D, Sobrero A et al. Cediranib with mFOLFOX6 versus bevacizumab with mFOLFOX6 as first-line treatment for patients with advanced colorectal cancer: a double-blind, randomized phase III study (HORIZON III). *J Clin Oncol* 2012; 30: 3588-3595.

Table 1 Relation between primary clinical endpoint, tolerance, and QoL outcomes results.

Clinical efficacy results for study primary endpoint and tolerance profiles	QoL results				
	Same	Better	Worse	Different	Total
	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>	<i>n</i>
Better clinical efficacy but same tolerance	1		1		2
Better clinical efficacy but worse tolerance	4	2			6
Better clinical efficacy but different tolerance	4	3			7
Same clinical efficacy and same tolerance	2			1	3
Same clinical efficacy but different tolerance	2	1			3
Same clinical efficacy but worse tolerance	2	2	3		7
Same clinical efficacy but better tolerance		1			1
Worse efficacy and worse tolerance		1		1	2
Total	15	10	4	2	31

“Better clinical efficacy” considered when a significant improvement was reported in primary endpoint

“Same clinical efficacy” indicates not significant results

Different indicates trials in which no global QoL differences were found, but the assessment of specific symptoms favored one or more of the arms

Appendix Supplementary Table SA1 Major first and second-line phase III colorectal cancer trials with health-related quality of life as an endpoint.

Study§	First author	N [¶]	Study arms	Endpoints			QoL instruments	Longitudinal statistical analysis	Time frame QoL measurement	Effect			
				Primary	Secondary	Tertiary				Toxicity	QoL	OS	PFS
	de Gramont ⁴⁵	422 (351 evaluated)	5FU2LV v FOLFOX	PFS	RR, OS, QoL, toxicity		EORTC QLQ-C30	TTD	At baseline and every fourth treatment cycle	More grade 3/4 neutropenia, grade 3/4 diarrhea, and grade 3 neurosensory toxicity on FOLFOX	Not significantly different; At cycle 4&8, emotional functioning improved on both arms, nausea or vomiting worsened on FOLFOX; TTD in global health status prolonged in the FOLFOX arm	Improved but not significantly; Survival without PD or deterioration in global health status was longer in patients treated with FOLFOX	Significantly improved in the FOLFOX arm
AVF2107g	Hurwitz ⁴⁸ Kabbinavar ⁶⁹	813	IFL v IFL + Bev	OS	PFS, ORR, DR, QoL		FACT-C total, CCS, TOI-C	TTD	At baseline, at 6 wk for the first 24 wk and q 12 wk thereafter	More grade 3-4 AE (specifically hypertension) in the IFL + Bev arm	Not significantly different	Significantly improved in the IFL + Bev arm	Significantly improved in the IFL + Bev arm
AGITG MAX	Tebbutt ⁷⁰ Stockler ⁷¹	265	Cap v Cap + Bev v Cap + Bev + MMC	PFS	OS, RR, toxicity, QoL		EQ-5D, EORTC QLQ-C30, CTAQ	Regression methods that accounted for repeated measures (i.e., generalized estimating equations with a compound symmetric correlation structure)	At baseline, at 3, 6, 9, 12 wk and q 6 wk until progression	Not significantly different; Significantly more HFS in Cap + Bev	Not significantly different	Not significantly different	Significantly improved in the Cap + Bev ± MMC arms
CAIRO3	Koopman ⁷²⁻⁷³ Punta ⁷⁴ Simkens ⁷⁵	558 (492 evaluated)	Cap + Bev v observation	PFS2	OS, QoL		EORTC QLQ-C30	GLMM adjusted for baseline and stratification factors	Before randomization and q 9 wk thereafter	More grade 3-4 AE (specifically grade 3 HFS) in the Cap + Bev arm	Not significantly different	Significantly extended (adjusted OS) with maintenance Cap +	Significantly extended with maintenance Cap + Bev

Formatted: Font: Italic

												Bev	
CORRECT	Grothey ⁷⁶	1052	Regorafenib v Placebo	OS	PFS, ORR, DCR, safety	DR/SD, QoL and health utility values	EORTC QLQ-C30, EQ-5D	TTD		More grade 3-4 AE in the regorafenib arm	Not significantly different	Significantly improved in the regorafenib arm	Significantly improved in the regorafenib arm
EPIC	Sobrero ⁷⁷	1298	Iri + Cetux v Iri	OS	PFS, RR, QoL		EORTC QLQ C30	Wei-Lachin test of multiple components	At 3 wk and q 6 wk thereafter	Toxicity profile similar but increased acneiform rash and diarrhea in the Iri + Cetux arm	Significantly improved in the Cetux + Iri arm,	Not significantly different	Significantly improved in the Cetux + Iri arm
CALGB/SWG 80405	Naughton ⁷⁸ Venook ⁷⁹	1137	FOLFOX or FOLFIRI + Bev or Cetux or Bev + Cetux	OS	PFS, QoL		EORTC QLQ-30, DSQ Scale		Not defined	No new toxicities emerged in this study. With bevacizumab, common side effects were hypertension, headache, mucositis, nosebleed, diarrhea, rectal bleeding, loss of appetite, fatigue, and weakness, while cetuximab patients were more likely to have acneiform rash, pruritus, changes in fingernails and toenails, infections, fatigue, and low serum electrolyte levels. FOLFOX was associated with more neuropathy, whereas FOLFIRI caused more alopecia and diarrhea.	Not significantly different in global health functioning or other items/subscales; Significant differences were found across arms in skin symptoms, limitations in social activities due to skin condition, and concerns about appearance; fewer social limitations and appearance concerns with Bev than with Cetux alone or Cetux + Bev	Not significantly different	Not significantly different
CAIRO2	Tol ⁸⁰	755	Cap + Ox +	PFS	OS, RR,		EORTC	Wilcoxon	At baseline (2	More grade 3 or	Overall and global not	Not	Significantly

			Bev (CB) v Cap + Ox + Bev + Cetux (CBC)		QoL	QLQ-30	rank-sum test	wk before randomizatio n) and q 9 wk thereafter until the end of study treatment	4 AE (attributed to Cetux) in the CBC arm	significantly different at baseline, after treatment significantly improved in the CB arm	significantl y different	reduced PFS in the CBC arm
NSAB Project Trial C-06	Lembersky ⁸¹ Kopec ⁸²	1608	oral UFT + LV v Intravenous FU/LV	DFS, OS	QoL	FACT-C, QLRS, SF- 36, SDS, SCL	GLMM	At baseline and at 1 year	Grade 3-4 AE similar between arms	Not statistically different for FACT-C, within the subscales of colon-specific, physical, emotional, social, and functional health, or QLRS scores; Statistically significant but small differences for SF-36 favoring FU/LV, for SDS & SCL favoring UFT UFT_LV associated with improved convenience of care	Not significantl y different	Not significantly different (and DFS)
CALGB 9481	Kemeny ⁸³	135	HAI v systemic bolus FU/LV	OS	RR, recurrence, toxicity, QoL, cost, the influence of molecular markers	SF-36, MSAS; MOS, MOS-SFQ	PMM	Before treatment and q 3 mo during treatment until mo 18	More grade 3 neutropenia and stomatitis in the HAI arm and higher bilirubin level in the FU/LV arm More men than women in the HAI arm had biliary toxicity	Significantly improved physical functioning in the HAI arm at 3 and 6 mo; No differences in social functioning, role functioning-emotional, or general health perceptions; At 12 mo, a level of physical functioning much higher in the HAI arm	Significantl y longer in the HAI arm	Significantly longer RR in the HAI arm
AIO KKK 0207	Hegewisch- Becker ⁸⁴ Arnold ⁸⁵ Quidde ⁴⁰	837	FP +Bev v Bev v no further treatment	TFS	OS, PFS1, QoL, safety	EORTC QLQ-CR29, EORTC QLQ-C30, and other instruments	Not defined	At screening, and then q 6 wk during induction and maintenance phase to a minimum of 24 wk after randomizatio n	Grade 3 AE low and similar in all arms	Not significantly different	Not significantl y different	Significantly different in the PF + Bev and Bev arms No significant difference in TFS overall
MRC CR06	Maughan ⁸⁶	905	Raltitrexed vs FU/LV vs FU	OS	QoL, toxicity, RR, cost,	HADS and EORTC QLQ-C30	Compared by standardizatio n of the	At baseline, and q 6 wk thereafter	More toxic effect in the raltitrexed arm	Worse for emotional functioning in the raltitrexed arm, HADS	Not significantl y different	Not significantly different

					acceptability of treatment to patients	and six trial- specific questions	subscale score: 0–24= nil, 25– 49=mild, 50– 74=moderate, 75– 100=severe			reduction in anxiety in FU and raltitrexed;		
CRYSTAL	Van Cutsem ^{87,88} Lang ⁸⁹	1198 (672 evaluated)	FOLFIRI + Cetux v FOLFIRI alone	PFS	OS, RR, DCR, QoL, safety	EORTC QLQ-C30 and social functioning scales	PMM	At randomization, q 8 wk before each cycle and at final tumor assessment	A minor increase in grade 3 and 4 diarrhea was seen with the addition of Cetux	Neither significantly improved nor negatively impacted	Significantly improved in the FOLFIRI + Cetux arm in <i>KRAS</i> WT patients	Significantly improved in the FOLFIRI + Cetux arm in <i>KRAS</i> WT patients; Significantly higher RR in patients with or without symptoms at baseline, and with improved and earlier symptom relief in those whose tumors had responded
NORDIC VII trial	Tveit ⁹²	571 (220 evaluated)	Continuous FLOX v Continuous FLOX + Cetux v intermittent FLOX + Cetux	PFS	OS, ORR, R0 resection rate, safety, QoL	Not defined	Not defined	Not defined	Grade 3 to 4 neutropenia similar in all arms; More grade 3 skin toxicity/rash in patients treated with Cetux	Not significantly different	Not significantl y different	Not significantly different
COIN	Adams ⁵¹	1630	Continuous and fluoropyrimidi ne +/- Cetux or Intermittent chemotherapy	OS	FFS, RR, toxicity, QoL, cost- effectiveness	EORTC QLQ C30 and five other questions at baseline	Ordinal logistic regression odds model	At baseline, at 6 and 12 wk, and q 12 wk thereafter	More grade 3 or worse hematological toxic effects with continuous treatment, while nausea and vomiting more common on intermittent treatment; More grade 3 or worse peripheral	Not different at baseline; at 24 wk, improvement in the intermittent arm (significant benefits in role and social functioning as well as for several symptom scales)	Not significantl y different	Not significantly different (the same for FFS)

									neuropathy and HFS in continuous treatment			
PRIME	Douillard ^{49,50}	1183	Pan + FOLFOX4 v FOLFOX4	PFS	OS, ORR, QoL, safety	EQ-5D HIS/OHR	MMRM	q 4 wk (± 1 wk) on study treatment and at the safety follow-up visit, and q 8 wk after the safety visit but before disease progression	Adverse event rates were generally comparable across arms with the exception of toxicities known to be associated with anti-EGFR therapy. More skin toxicity in Pan-FOLFOX4	No significantly different; deterioration of performance status occurred later in the irinotecan combination than in those treated with 5FU/LV alone	Significantly improved in the Pan + FOLFOX4 arm in WT KRAS patients and in patients who develop skin toxicity grade 2–4	Significantly improved (and ORR) in the panitumumab + FOLFOX4 arm in WT KRAS patients and in patients who develop skin toxicity grade 2–4
GONO group trial	Falcone ⁵⁰	244	FOLFOXIRI vs FOLFIRI	RR	OS, PFS, post-chemotherapy R0 resection, QoL, safety	EORTC QLQ-C30	Not defined	q 2 wk during treatment and q 8 wk during follow-up (36 mo)	More grade 2 to 3 peripheral neurotoxicity and grade 3 to 4 neutropenia in the FOLFOXIRI arm; Febrile neutropenia and grade 3 to 4 diarrhea not significantly different between arms	Not significantly different;	Significantly improved in the FOLFOXIRI arm	Significantly improved (and RR) in the FOLFOXIRI arm
NCIC CTG & AGITG CO. 17	Au ⁵¹ Jonker ⁹²	572	Cetux + BSC v BSC	OS	PFS, RR, QoL, safety	EORTC QLQ-C30	PMM + LMM	At 4, 8, 16, 24 wk	Not significantly different in the incidence of grade 3 or higher AE, with the exception of more rash of any grade, hypomagnesemia of any grade, infusion reactions of any grade, infection without neutropenia, confusion, and	Significantly improved in GHS and PF at 16 wk in the BSC arm in <i>KRAS</i> WT patients	Significantly improved in the Cetux + BSC arm	Significantly improved in the Cetux + BSC arm

									pain in the Cetux arm; A rash of grade 2 or higher was strongly associated with improved survival			
NCIC CTG & AGITG CO. 20	Righash ⁶⁶ Siu ⁹³	750	Cetux + Briv v Cetux + Placebo	OS	PFS, ORR, DR, QoL, safety	EORTC QLQ-C30	TTD	At 2, 4, 6, 8, 12, 16, 24 wk	Statistically significantly higher incidence of grade 3 non-hematologic AE: fatigue, hypertension, rash, and gastrointestinal toxicities including abdominal pain, diarrhea, dehydration, and anorexia in the Cetux + Briv arm; In general, much higher rates of certain AE (e.g., fatigue, diarrhea) were reported by patients than were captured by AE reporting	Cetux + Briv worsened time to QoL deterioration for <i>KRAS</i> WT patients	Not significantl y improved	Significantly improved (and ORR) in the Cetux + Briv arm
Study 181*	Peeters ^{94,95}	1186	Pan + FOLFIRI v FOLFIRI alone	PFS and OS (co-primary end points by tumor <i>KRAS</i> status)	ORR, QoL	EQ-5D HIS/OHR	MMRM	q 8 wk until PD	AE rates generally comparable with the exception of known toxicities associated with anti-EGFR therapy (more grade 3 to 4 AE in the Pan + FOLFIRI arm)	No significant difference for the EQ-5D HIS and EQ5-D OHR	Significantl y improved in the Pan-FOLFIRI arm in <i>KRAS</i> WT patients	Significantly improved (and ORR) in the panitumuma b-FOLFIRI arm in <i>KRAS</i> WT patients
PICCOLO*	Seymour ⁹⁶	597	Iri + Pan v Iri	OS	PFS, RECIST response, QoL,	EORTC QLQ-C30, EQ-5D, DLQI	Not defined	At baseline and at 12 and 24 wk	More grade 3 or higher, diarrhea, lethargy, skin toxicity,	At 24 wk, EORTC QLQ-C30 significantly better in the Iri + Pan arm, but QoL symptom scores	No significantl y different	Significantly different (and RECIST

					toxicity				infection, and neutropenia in the Iri + Pan arm as were any hematological, any non-hematological, or a grade 3 or higher toxicity of any type	worsened with Ir + Pan; No data for DLQI		reponse) in the Iri + Pan arm
X-ACT	Twelves ⁹⁷	1987	Oral Cap v bolus FU/LV	DFS	RFS, OS, 3-year DFS, safety, QoL	EORTC QLQ-C30	Not defined	At baseline, and at 7, 16, 25 wk in the Cap arm; and at 9, 17, 25 wk in the FU/LV arm	Significantly less grade 3 or 4 toxic effects in the Cap arm	At 25 wk, significant increase from baseline (<5% in raw scores)	Equivalent in both arms	DFS equivalent in both arms, RFS improved in the Cap arm
EORTC study 40952	Kohne ⁹⁸	497	q.wk. 24h FU ± LV v bolus FU/LV	OS	PFS, RR, safety, QoL	EORTC QLQ-C30	Pearson's correlation matrix	At baseline	More toxicity in the 24hFU + LV arm	No significantly different	No significantl y different	Significantly longer for 24hFU + LV vs two other arms, RR not significant
Irinotecan Study Group	Saltz ⁹⁹	1400	FOLFIRI v FU/LV v Iri alone	PFS	ORR, OS, safety, TTP	EORTC QLQ-C30	Analysis of variance for repeated measures	At start of each treatment cycle	More grade 3 (severe) diarrhea in the FOLFIRI arm; but incidence of grade 4 (life-threatening) diarrhea was similar in both arms; Less grade 3 or 4 mucositis, grade 4 neutropenia, and neutropenic fever in the FOLFIRI arms. Adding Iri to FU/LV did not compromise the QoL	No significantly different	Significantl y improved in the FOLFIRI arm	Significantly improved (and RR) in the FOLFIRI arm
	Conroy ¹⁰⁰ Ducreux ¹⁰¹	306	XELOX v FOLFOX6	ORR	OS, PFS, TTR, RD, QoL	EORTC QLQ-C30, FACIT-CCSQ	MMRM	At baseline, and at Cycle 3 and 6 visits (XELOX) or Cycle 4 and 8	Significantly more grade 3/4 thrombocytopenia and diarrhea in the XELOX arm,	No significantly different	No significantl y different	Not significantly different; Significantly improved for

								visits (FOLFOX-6) and final visit (Day 169)	but significantly less grade 3/4			ORR, in the FOLFOX6 arm
MRC FOCUS2	Seymour ¹⁰²	459	FU + Levofolinate v Ox + FU v Cap v Ox + Cap	PFS + QoL	RR, OS, toxicity	EORTC QLQ-C30	Mann- Whitney test	At baseline (117-item CHA) and at 12 wk (OTU)	Increased risk of having a grade 3 or worse event during the first 12 wk in the Ox + Cap arm. The risk of having any grade 3 or worse toxic eff ect was not significantly increased with Ox, but was higher with Cap (specifically nausea, vomiting, diarrhea, anorexia, and HFS) than with FU	No significantly different; fewer baseline symptoms were predictive of better OTU	No significantl y different	No significantly different
	Haller ^{103*}	628	Iri + Ox v Iri alone	OS	ORR, DR, TTP, QoL (TTSW), toxicity	TRSQ (analgesic consumptio n, pain intensity, weight, KPS)	Two-sided unstratified log-rank test; Association of improvement in TRS and ORR by Spearman correlation coefficient	At baseline (within 7 days before randomizatio n and the first dose of study drug) and q 3 wk (before chemotherapy for up to 3 mo after the last dose of study medication	Grade 3 to 4 toxicities were comparable between arms; more granulocytopenia , diarrhea, and sensory disturbances in the Iri + Ox arm	Significantly improved TTSW (TRS, specifically analgesic consumption, pain) in the Iri + Ox arm	Significantl y improved in the Iri + Ox arm	Significantly improved ORR and TTP in the Iri + Ox arm

ASPECCT	Price ¹⁰⁴	999	Pan v Cetux	OS	PFS, ORR, DR, TTR, TTF, QoL, toxicity	EQ-5D HIS/VAS, FACT FCSI	MMRM	From study day 1 to the last day of treatment or PD, up to wk 85	Any grade and grade 3–4 AE similar between arms; Grade 3–4 infusion reactions lower and grade 3–4 hypomagnesaemi a higher with Pan	No significantly different	No significantl y different	No significantly different
	Van Cutsem ¹⁰⁵ Siena ¹¹	463 (391 evaluated)	Pan + BSC v BSC alone	PFS	ORR, OS, PROs	FACT FCSI, EORTC QLQ-C30, EQ-5D HIS/VAS	LVCf and a slope method	At baseline and at q 2 wk or monthly during treatment, and at the 30-day safety follow- up visit	Skin toxicities, hypomagnesaemi a, and diarrhea more common with Pan	At wk 8, lack of PD was associated with significantly and clinically meaningful lower CRC symptomatology for both arms and higher HRQoL for Pan patients only. Lack of PD was associated with better symptom control, HRQoL, and OS	No significantl y different	Significantly improved (and ORR) in the Pan + BSC arm
HORIZON III	Schmoll ¹⁰⁶	709 713	mFOLFOX6 + Cediranib v mFOLFOX6 + Bev	PFS	OS, ORR, DR, safety, tolerability, resection rate, HRQoL	FACT-C (FACT-G + CCS)	Stratified log- rank test	At ≤ 4 wk before the start of study treatment; at 8-wk intervals until wk 24, and every 12 wk until disease progression or death	Grade ≥ 3 neutropenia and diarrhea, grade ≥ 3 AE, serious AE, and AEs leading to permanent discontinuation of drug more common in the Cediranib arm.	Time to worsening of symptoms significantly shorter in the Cediranib arm; Statistically significant intertreatment differences favoring Bev for TOI, total FACT Score, and CCS. No difference between progressive/nonprogressi ve, patients and responders/ nonresponders	No significantl y different	No significantly different

*2nd line

§ PubMed and abstracts of major conferences (ASCO, ASCO GI, and ESMO) proceedings were searched for articles on phase first and second-line phase III clinical trials including health-related quality of life as an outcome measure using terms “quality of life”, “endpoint”, “phase III”, “colorectal”, “first-line “, “second-line “, and “randomized” and with no restrictions on publication date.

Abbreviations: EORTC QLQ C30, European Organization for Research and Treatment of Cancer Quality of Life Questionnaire C-30; FACT-C, Functional Assessment of Cancer Therapy–Colorectal; CCS, Colorectal Cancer Subscale; TOI, Trial Outcome Index, EQ-5D, EuroQoL-5 Dimension Questionnaire; SF-36, Short Form-36 questionnaire; QLRS, Quality of Life Rating Scale; HADS, Hospital Anxiety and Depression Scale; BSC, best supportive care; HAI, hepatic arterial infusion; THP, time to hepatic progression; MOS, Medical Outcomes Study; MOS-SS, Medical Outcomes Study Social Support Questionnaire; MOS-SFQ, Medical Outcomes Study Sexual Functioning Scale; MSAS, Memorial Symptom Assessment Scale; HIS, health state index; VAS, visual analogue scale; DLQI, Dermatology Life Quality Index; SDS, Symptom Distress Scale; SCL, Treatment-specific Symptom Checklist; CTAQ, Chemotherapy Acceptability Questionnaire; QLRS, Quality of Life Rating Scale; DSQI, Dermatology-Specific Quality of Life Scale; OHR, overall health rating; GHS, global health

status; OHS, overall health symptoms; FACIT-CCSQ, Functional Assessment of Chronic Illness Therapy Chemotherapy Convenience and Satisfaction Questionnaire; FCSI, FACT Colorectal Symptom Index; CHA, comprehensive health assessment; OUT, overall treatment utility (the viewpoint of both patient and clinician) FU/LV, fluorouracil; leucovorin; Cetux, Cetuximab; Cap, capecitabine; Iri, irinotecan; Ox, oxaliplatin, Bev Bevacizumab; FOLFOX, fluorouracil; leucovorin plus oxaliplatin; FOLFIRI, fluorouracil; leucovorin plus irinotecan; FOLFOXIRI, fluorouracil; leucovorin, oxaliplatin plus irinotecan; XELOX, capecitabine plus oxaliplatin; chronoFLO4, fluorouracil, leucovorin plus oxaliplatin for 4 days; FOLFOX, fluorouracil; leucovorin plus oxaliplatin; IFL, irinotecan, fluorouracil plus leucovorin; FP, fluoropyrimidine; MMC, mitomycin C; Pan, panitumumab; UFT, tegafur plus uracil; FLOX, fluorouracil, folinic acid plus oxaliplatin; FLIRI, fluorouracil, folinic acid plus irinotecan; BRIV, cetuximab plus brivanib alaninate; QoL, quality of life; OS, overall survival; PFS, progression-free survival; RR, response rate; ORR, overall response rate; TTP, time to tumor progression; GLMM, general linear mixed-effects model, MMRM, mixed-effect model for repeated measures; PMM, pattern mixture models; LMM, linear mixed model; LVCF, last value carried forward; TTD, time to deterioration; PFS1, time to first progression; TTR, time to response; DR, duration of response; DFS, disease-free survival; SD, stable disease; DCR, disease control rate; PD, disease progression; TFS, time to failure strategy; RFS relapse-free survival; FFS, failure-free survival; TTSW, time to tumor-related symptomatic worsening; TRSQ, tumor-related symptom questionnaire; KPS, Karnofsky performance status; HFS, hand-foot syndrome; CT, chemotherapy; q, every; wk, week; q.wk., weekly; mo, month; WT, wild-type; AE, adverse events

Appendix Supplementary Table SA2. Compliance rate for HRQoL questionnaires in adjuvant and metastatic colorectal cancer studies

Study	Author	Compliance rate (completed/expected) at different time points			
		At baseline	During chemotherapy		After treatment
NSAB C-06	Kopec ⁸²	93% FU 94% UFT		73% FU 79% UFT	1-year 62% FU 67% UFT
MRC CR06	Maughan ⁸⁶	84% in total	6 wk 12 wk	64% in total 62% in total (43% of those alive at 12 wk completed questionnaires at three assessment points)	
COIN	Adams ⁵¹	67% in total	6/12 wk	60% continuous 55% intermittent	
AGITG MAX	Tebbutt ⁷⁰	91% Cap 92% Cap + Bev 91% Cap+ Bev + MC			
NCIC CTG & AGITG CO.17	Au ⁹¹	94% in total	16 wk	67% Cetux + BSC 47% BSC	
NCIC CTG & AGITG CO.20	Ringash ⁶⁶	95% Cetux + Briv 97% Cetux + Placebo	2 wk	80% Cetux + Briv 87% Cetux + Placebo	
CRYSTAL	Van Cutsem ^{87,88} Lang ⁸⁹	75% in each arm	4 wk	80% Cetux + Briv 85% Cetux + Placebo	
			6 wk	71% Cetux + Briv 78% Cetux + Placebo	
			8 wk	69% Cetux + Briv 77% Cetux + Placebo	
			12 wk	72% Cetux + Briv 69% Cetux + Placebo	
			16 wk	68% Cetux + Briv 62% Cetux + Placebo	
			24 wk	59% Cetux + Briv 47% Cetux + Placebo	
			8 wk	74% FOLFIRI 79% FOLFIRI + Cetux	
			16 wk	69% FOLFIRI 68% FOLFIRI + Cetux	
			24 wk	66% FOLFIRI 67% FOLFIRI + Cetux	

			32 wk	57% FOLFIRI 51% FOLFIRI + Cetux
			40 wk	56% FOLFIRI 61% FOLFIRI + Cetux
			48 wk	53% FOLFIRI 60% FOLFIRI + Cetux
			56 wk	46% FOLFIRI 57% FOLFIRI + Cetux
			64 wk	70% FOLFIRI 63% FOLFIRI + Cetux
			Final visit	35% FOLFIRI 30% FOLFIRI + Cetux
EPIC	Sobrero ⁷⁷		15 wk	56.3% Cetux + Iri 56.1% Iri
AIO KRK 0207	Hegewisch- Becker ⁸⁴ Arnold ⁸⁵ Quidde ⁴⁰	89% in total	6 wk	99%
			12 wk	98%
			18 wk	97%
			24 wk	97%
	Conroy ¹⁰⁰ Ducreux ¹⁰¹	97.6% XELOX 97.5% FOLFOX6	Cycle 3/4 61.9%	77.8% XELOX 58.0% FOLFOX6
			Cycle 6/8	61.9% XELOX 61.3% FOLFOX6
			Final visit	50% XELOX 54.6% FOLFOX6
AVF2107g ‡	Hurwitz ⁴⁸ Kabbinavar ⁶⁹ ‡	30.4% IFL + Placebo [†] 30.2% IFL + Placebo ^{††} 30.9% IFL + Placebo ^{†††} 30.3% IFL + Bev [†] 30.1% IFL + Bev ^{††} 30.3% IFL + Bev ^{†††}		
PRIME	Douillard ^{149,50}			57% in WT <i>KRAS</i> patients 45% of patients had missing data (similar proportions for both treatment arms)

Van Cutsem ⁸⁷ Siena ¹¹	4 wk	81.8% Pan + BSC* 55.6% BSC 82.2% Pan + BSC** 56.0% BSC
	8 wk	48% Pan + BSC* 20.2% BSC 48.5% Pan + BSC** 20.7% BSC
	12 wk	39.4% Pan + BSC* 6.1% BSC 39.0% Pan + BSC** 6.0 % BSC
	16 wk	26.8% Pan + BSC* 3.0 % BSC 26.8% Pan + BSC** 3.0% BSC
HORIZON III	Schmoll ¹⁰⁶	84% to 98% (Compliance with the FACT-C questionnaire at all time points)

*EQ-5D questionnaire

**FCSI questionnaire

†TOI-C questionnaire

††FACT-C questionnaire

†††CCS questionnaire

‡The baseline and at least one post-baseline assessment

Abbreviations: FU, fluorouracil; UFT, tegafur plus uracil; Cap, capecitabine; Bev, bevacizumab; MMC, mitomycin C; Cetux, Cetuximab; BRIV, cetuximab plus brivanib alaninate; BSC, best supportive care; FOLFIRI, fluorouracil; leucovorin plus irinotecan; XELOX, capecitabine plus oxaliplatin; FOLFOX, fluorouracil; leucovorin plus oxaliplatin; IFL, irinotecan, fluorouracil plus leucovorin; Pan, panitumumab; GQoL, Global quality of life; EQ-5D, EuroQoL-5 Dimension Questionnaire; wk, week