Article:
Fokas, E, Glynne-Jones, R, Appelt, A orcid.org/0000-0003-2792-9218 et al. (13 more authors) (2020) Outcome measures in multimodal rectal cancer trials. Lancet Oncology, 21 (5). e252-e264. ISSN 1470-2045

https://doi.org/10.1016/S1470-2045(20)30024-3

Outcome measures in multimodal rectal cancer trials: Which endpoints should we use?

Prof Emmanouil Fokas MD DPhil[1, 2, 3, #], Dr Robert Glynne-Jones MRCP, FRCR[4, #], Dr Ane Appelt PhD[5], Prof Regina Beets-Tan MD PhD[6,7], Prof Geerard Beets MD PhD[7,8], Prof Karin Haustermans MD PhD[9], Prof Corrie Marijnen MD[10], Prof Bruce D. Minsky MD[11], Dr Ethan Ludmir MD[11], Prof Phil Quirke FRCPath FMedSci PhD[12], Prof David Sebag-Montefiore MRCP, FRCR[5], Prof Julio Garcia-Aguilar MD, PhD[13], Prof Maria Antonietta Gambacorta MD, PhD[14], Prof Vincenzo Valentini MD[14], Prof Marc Buyse ScD[15,16,§], Prof Claus Rödel MD[1, 2, 3,§]

[1]Department of Radiotherapy of Oncology, University of Frankfurt, Germany
[2]German Cancer Research Center (DKFZ), Heidelberg, Germany; German Cancer Consortium (DKTK), Partner Site: Frankfurt
[3]Frankfurt Cancer Institute (FCI)
[4]Department of Radiotherapy, Mount Vernon Centre for Cancer Treatment, Northwood, Middlesex, UK
[5]Institute of Medical Research at St James's, University of Leeds, Leeds, UK
[6]Department of Radiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands
[7]GROW School for Oncology and Developmental Biology, Maastricht University, The Netherlands
[8]Department of Surgery, Netherlands Cancer Institute Amsterdam, Amsterdam, The Netherlands
[9]Department of Radiation Oncology, University Hospital Leuven, Leuven, Belgium
[10]Department of Radiation Oncology, Netherlands Cancer Institute Amsterdam, Amsterdam, The Netherlands
[11]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[12]Division of Pathology and Data Analytics, School of Medicine, Leeds University, Leeds, UK
[13]Colorectal Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York
[14]Department Radiation Oncology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Roma, Italy
[15]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium
[16]International Drug Development Institute, San Francisco, CA, USA

#joint first authors; §joint last authors

Address for Correspondence
Prof Emmanouil Fokas, MD DPhil
Department of Radiotherapy and Oncology,
University of Frankfurt
Theodor-Stern-Kai 7
60590 Frankfurt, Germany
Phone: +49-(0)69-63015130
Fax: +49-(0)69-63015091
E-Mail: emmanouil.fokas@kgu.de

SUMMARY

There has been a large variability regarding the definition and choice of primary endpoints in phase 2 and phase 3 multimodal rectal cancer trials. This has resulted in inconsistency and difficulty in data interpretation. Also, surrogate properties of early and intermediate endpoints have not been systematically assessed. We provide a comprehensive review of clinical and surrogate endpoints used in trials for non-metastatic rectal cancer. The applicability, advantages and disadvantages of these endpoints are summarized, and recommendations on the clinical endpoints for the different phase trials are provided, including the context of limited- or non-operative management for organ preservation. Finally, we discuss how early and intermediate endpoints, including patient-reported outcomes and involvement of patients in decision making, can be used to guide trial design and facilitate consistency in reporting trial results in rectal cancer.

INTRODUCTION

Rectal cancer is among the best examples in oncology of how progress is derived from multimodal treatment. Preoperative 5-fluorouracil-(5-FU)-based chemoradiotherapy (CRT) or short-course radiotherapy (SCRT) followed by total mesorectal excision (TME) have substantially reduced locoregional recurrence (LRR). However, most trials have failed to demonstrate an improvement in disease-free survival (DFS) or overall survival (OS)[1,2].

In phase 3 cancer trials, OS is the most objectively defined endpoint and has, for a long time, been considered the standard measure of treatment efficacy. OS requires a large sample size and long-term follow-up, is costly and carries the risk of the investigated treatment losing novelty by the time of trial completion. Moreover, OS can be confounded by effective successive treatment lines in case of disease progression or recurrence, and competing risks for non-cancer-related deaths, especially among older and frail patients, may dilute the observed treatment effect[3]. As such, there has been an increased interest in early or intermediate endpoints as surrogate measures of long-term endpoints in oncology trials[4].

As outcome measures have been derived at several different time points during and after multimodal treatment, rectal cancer may be ideally suited to address the complex interplay between early, intermediate and long-term endpoints (**Figure 1**). Here, we review surrogate and time-to-event clinical endpoints used in non-metastatic rectal cancer in the context of completed and ongoing randomised multimodal clinical trials, including those of limited- or non-operative management for organ preservation. The applicability, advantages and disadvantages of endpoints are summarised. The review focuses mostly on the neoadjuvant setting. The impact of rectal cancer treatment on quality of life (QoL) has been increasingly recognized and might be as important for some patients, thus we also outline the evidence on patient-reported outcome measures (PROMs) in rectal cancer. Finally, we discuss how early and intermediate endpoints can potentially be used to guide trial design and facilitate consistency in reporting trial results in rectal cancer.

METHODS

Literature search and selection criteria

We performed a computerised literature search using Pubmed, MEDLINE, Web of Science and the Cochrane Library for full published articles and abstracts from international meetings from January 1993 to September 2019 supplemented by hand searching of abstracts from recent international meetings. MeSH terms or combined free terms included “rectal cancer”, “clinical

endpoint", "surrogate", "surrogate endpoint", "composite endpoint", "randomised trials", "phase 2/3" (or phase II/III)", "survival", "radiotherapy", "short-course radiotherapy", "chemoradiation", "preoperative", "neoadjuvant", "organ preservation", "non-operative management", "Quality of life (QoL)" and "PROMs". Only papers published in English were reviewed. For the purpose of the present review, only randomised phase 2 and phase 3 clinical trials with clearly defined clinical endpoints were selected, while single arm, prospective non-randomised trials and studies that lacked of clear clinical endpoint definition were excluded. We also excluded titles if considered irrelevant to the scope of this review or duplicate publications. Unpublished ongoing randomised phase 2 and phase 3 studies were cited using either the NCT trial number (https://clinicaltrials.gov), the ISRCTN registry (https://www.isrctn.com) or EudraCT number (https://eudract.ema.europa.eu). Of the 625 full-text articles or abstracts assessed for eligibility, 71 trial publications were relevant. The final reference list was generated on the basis of relevance to the broad scope of the present review.

RESULTS

Early endpoints to reflect tumour response at surgery – the concept of surrogacy

Tumour response to CRT/SCRT varies considerably in rectal cancer. Early variables to assess tumour response, such as downstaging, pathological complete response (pCR), tumour regression grading (TRG), the neoadjuvant rectal score (NAR), completeness of local resection (R0/1/2) and circumferential resection margin (CRM) have been proposed to reflect treatment efficacy and patients' prognosis, and have been suggested as potential surrogates for longer term outcomes, such as LRR, DFS, and OS (**Figure 1**).

Tumour response is a dynamic process associated with tumour-related factors, such as size, histology, and molecular profile, and with treatment-related factors, such as RT dose and fractionation, combination with chemotherapy, and the time interval between neoadjuvant treatment and surgery. The complex interplay between these factors and the varying time points of response assessment makes interpretation of early efficacy endpoints challenging. Although increased tumour response may predict favourable outcome on the level of individual patients within a given treatment, this does not necessarily imply superior outcomes in comparative trials. For example, the Polish I and Tran-Tasman 01.04 randomised trials of SCRT followed by immediate surgery versus 5-FU-CRT followed by delayed surgery showed significantly increased pathologic downstaging and pCR in the 5-FU-CRT groups, yet, no improvement in

LRR, DFS or OS[5,6]. Thus, differences in early efficacy endpoints induced by different treatments may not be valid surrogates for long-term oncological outcomes.

The basic requirement for surrogacy is that changes of the surrogate between treatment interventions should translate to changes of the true clinical endpoints. The establishment of surrogate endpoints is challenging as it requires rigorous statistical validation using large trial datasets. Prentice proposed four criteria (PC1-4) to validate surrogacy[7]. First, the treatment must significantly impact on the true clinical endpoint (PC 1). Second, the treatment must have a significant effect on the surrogate (PC 2). Third, the surrogate must have a statistically significant effect on the true endpoint (PC 3) and fourth, the full impact of treatment on the true endpoint should be captured by the surrogate (PC 4). However, PC4 is characterised by inherent difficulties and methodological limitations in its interpretation. An alternative validation method has been established by Buyse et al. based on the "correlation approach" that involves demonstration of surrogacy at two levels[4,8]. First, "individual-level" surrogacy reflects the correlation of the surrogate outcome with the clinical outcome, and can be demonstrated in a single trial dataset using standard correlation coefficients. Secondly, "trial-level" surrogacy is based on meta-analysis of phase 3 trials.

Downstaging, i.e. a decrease of the pathological versus the preoperative clinical T- and/or N category, has been used as an early endpoint in older trials to evaluate the efficacy of preoperative CRT/SCRT. Assessment of downstaging necessitates accurate baseline tumour T- and N-staging if it is to be considered valid. Magnetic resonance imaging (MRI) is now the standard diagnostic method in rectal cancer, but several older trials were using endoscopic ultrasound and/or CT imaging that were characterised by inherent methodological problems and could have led to either under- or overstaging. The CAO/ARO/AIO-94 trial confirmed that preoperative CRT induced significant downstaging compared with straight surgery. The subgroup of patients with ypT0N0 or ypT1-2N0 disease after preoperative CRT had excellent prognosis, however, those patients with remaining ypT3-4 and/or ypN1-2 after preoperative CRT had higher risk of developing distant metastases with poorer DFS compared to patients with corresponding pTNM tumour subgroups in the postoperative treatment arm. For the entire study cohort, DFS and OS remained identical in both arms, resembling the reverse of the "stage migration" or "Will Rogers phenomenon"[9]. Thus, downstaging may unmask different prognostic subgroups within a given treatment group, rather than reflect improved treatment efficacy between interventions. Also, there is no clear definition of downstaging that makes its use challenging. Some investigators used a reduction of either T or N categories by at least 1 level as definition, whereas others restricted this endpoint to a reduction of stage. As such,

validation of downstaging as a surrogate endpoint based on large phase 3 trial cohorts is currently lacking.

Pathological complete response (pCR) has constituted the primary endpoint in numerous early phase 1 and 2 trials, and some randomised phase 2 and 3 studies (**Tables 1-3**). Maas et al. demonstrated the strong correlation of pCR after CRT with better DFS and OS[10]. The excellent prognosis of patients with pCR could be, in part, explained by possible overstaging and overtreating of early-stage tumours. The pooled individual data analysis from the EORTC 22921 and FFCD 9203 phase 3 trials comparing the effect of preoperative RT with or without concurrent 5-FU showed significantly higher pCR rates and reduced 3-year LRR in the CRT versus RT arms, but no significant improvement in OS and poor trial-level surrogacy of pCR for OS[11]. Likewise, pCR was poorly correlated with 5-year OS – both at the individual and trial levels – in a meta-analysis of 22 randomised trials involving 10,050 patients treated with neoadjuvant RT/CRT[12]. Thus, intensification of local neoadjuvant treatment, e.g. by increasing RT dose or adding radiosensitzing agents, may improve pCR rates and potentially reduce LRR, however, the natural history of the disease, especially with respect to distant metastases may not be altered.

In the Lyon R90-01 randomised trial, a longer interval (6-8 versus 2 weeks) after completion of RT led to significant increase of patients with major pathologic response (pCR or few residual cells) without impact on LRR or OS[13]. Further extension beyond 8 weeks has been tested in the GRECCAR-6 (7 versus 11 weeks) and a British randomised trial (6 versus 12 weeks)[14,15]. While the first failed to show an increase of pCR with the prolonged interval (15% versus 17·4%, P=0·59), the second did (9% versus 20%, P<0·05). Similarly, improved pCR (18% vs 10%; P=0·027) was reported by Akgun et al. for an interval of >8 weeks vs <8 weeks after CRT[16]. Long-term outcomes of these trials are pending (**Table 1**). The available trial data are conflicting and do not suggest that simply increasing the interval beyond 6-8 weeks will result in better long-term outcomes. Also, the majority of the above studies were of small size that should be considered when interpreting the findings.

Several groups have used the prolonged interval between CRT/SCRT and surgery for adding neoadjuvant, rather than adjuvant chemotherapy, known as total neoadjuvant treatment (TNT). A prospective phase 2 cohort trial used preoperative CRT and sequentially increased the time point of surgery. Study group 1 underwent surgery 6 weeks after completion of CRT. Patients in study groups 2, 3 and 4 received two, four, or six cycles of FOLFOX during the waiting period before surgery performed 11, 15, and 19 weeks, respectively, after completion of CRT.

The pCR rate of patients treated in study group 1 was 18% compared with 25%, 30%, and 38%, respectively, for study groups 2-4[17]. Whether this strategy of both prolonging the RT-surgery interval and adding systemic therapy during this interval will lead to improved DFS and OS, and/or an increase in organ preservation, is currently being tested in the Organ Preservation in Rectal Adenocarcinoma phase 2 trial (OPRA; NCT02008656) and the RAPIDO phase 3 trial after SCRT (NCT01558921). The CAO/ARO/AIO-12 randomised phase 2 trial that compared the two TNT sequences, CT/CRT vs CRT/CT, demonstrated a pCR in 17% and 25%, respectively [18], however, long-term data are needed to show whether improved pCR translates into superior DFS (**Table 2**).

Lymph nodes status after preoperative CRT/SCRT constitutes an important prognostic factor for both local and distant recurrence[19]. In the CAO/ARO/AIO-04 trial, the 3-year cumulative incidence of LRR and distant recurrences for the entire study cohort were 2% and 10·6% for ypN0, 2·5% and 28·6% for ypN1, and 17·2%, and 48% for ypN2, respectively. Persistent lymph node metastases after neoadjuvant CRT reflect a highly aggressive tumour phenotype both resistant towards CRT and prone to distant metastases.

The NeoAdjuvant Rectal (NAR) score was proposed by the US NRG group to serve as a potential surrogate for OS in rectal cancer[20]. The NAR formula $[5\ ypN - 3(cT\text{-}ypT) + 12)^2/9.61]$ results in 24 distinct scores that range from 0 to 100. For ypT- and ypN-category, a relative weight of 3 and 5 has been suggested to reflect the impact of these variables, based on the nomogram of Valentini and colleagues[21]. The NAR score was classified as low (NAR<8), intermediate (NAR=8-16), and high (NAR>16) according to the tertiles of the scores of the randomised NSABP R-04 trial dataset. Lower NAR score was associated with better OS.

The individual-level surrogacy of NAR for DFS was demonstrated in the CAO/ARO/AIO-04 phase 3 trial using the Prentice criteria: the addition of oxaliplatin to preoperative 5-FU-CRT resulted in a significant DFS improvement (PC1), reflected in a significant shift towards lower NAR scores at surgery (PC 2). The NAR score was an independent predictor for DFS (PC3), and the treatment effect on DFS was captured by NAR (PC4)[22]. Yothers et al. showed a limited but significant association between change in NAR score and improvement in OS (r^2=0·13) by meta-analysing 5 randomised trials of neoadjuvant treatment; this trial-level association was nominally better than for pCR and OS (r^2=0·02)[23]. The prognostic value of NAR was also confirmed in the PAN-EX, a pooled analysis of the EXPERT and EXPERT-C phase 2 trials that tested induction chemotherapy followed by CRT in patients with high-risk rectal cancer[24].

NAR is currently used as primary endpoint in the NRG GI002 phase 2 trial platform (NCT02921256) for TNT to assess novel sensitizers (**Table 2**): The addition of veliparib, a PARP inhibitor, to standard CRT after induction FOLFOX chemotherapy failed to reach the primary endpoint of the trial, that is a 4-point reduction in NAR score[25].

As with other early endpoints, the NAR score has limitations. It is prone to the inherent limitations of baseline diagnostic staging as it takes into account the preoperative cT-category. In the large Dutch Cancer registry, the NAR score failed to a show a superior prognostic value compared to combined pT and pN for OS[26]. Also, it is difficult to interprete the clinical impact of small changes of the NAR score (e.g., a 4-point reduction) and to determine which numerical changes translate to improvements in LRR, DFS and OS. Moreover, surrogacy of NAR has been tested mainly for long-course CRT and may not be valid for other forms of preoperative interventions, such as SCRT or chemotherapy alone.

Tumour regression grading (TRG) is a semiquantitative assessment of residual tumour cells versus fibro-inflammatory tissue in the rectal wall after preoperative treatment. There is no consensus for a universally approved standardisation, and the studies using TRG were characterised by heterogeneity regarding methodological assessment, patient cohort, treatment modality, and the time interval between CRT and surgery[27]. Data from the CAO/ARO/AIO-94 and -04 studies suggest that TRG identified distinct prognostic groups independent of established prognostic factors such as the TNM classification system, and also fulfilled the PC1-4 for individual-level surrogacy[19,28] similarly to the NAR score[22]. Evidently, a shift of 24 scores, as in the case of NAR, or 3-5 tier semiquantitative TRG, may reflect treatment effects induced by different neoadjuvant regimens more accurately than binary endpoints, such as downstaging or pCR, and as such these treatment-induced changes have the potential of being better candidate surrogates for OS. More information on pathological assessment of pCR and TRG is shown in **Supplementary Material.**

The MRI-based TRG (mrTRG) was used as an endpoint in the MERCURY trial to stratify patients to poor (mrTRG 4-5) vs good response (mrTRG 1-3); a significantly worse DFS and OS was reported among patients with poor response[29]. Only a weak correlation between mrTRG and pathologic TRG was reported in the EXPERT and EXPERT-C phase 2 trials and, thus, further studies to assess the applicability of mrTRG as an endpoint are warranted[30]. The TRIGGER randomised feasibility trial (NCT02704520) is currently assessing the potential of MRI to guide treatment selection, including deferral of surgery, according to mrTRG after CRT.

The most important surgical margin is the **circumferential resection margin (CRM)** created around the mesorectum. Quirke and colleagues have shown that LRR is greatly increased and OS halved when tumour can be demonstrated at or within 1 mm from the radial surgical plane of resection. Involved CRM (≤1mm) was more common in patients with advanced stage, ulcerative growth pattern, poor differentiation, vascular invasion, poor TME quality, and abdominoperineal resection. The prognostic value of the CRM for LRR was even higher after neoadjuvant CRT/SCRT than when no preoperative therapy has been applied, and CRM is also a powerful predictor of distant metastases and OS[31].

Despite their strong prognostic relevance, R-status and CRM have only been used as secondary endpoints in phase 3 trials. The randomised Polish II trial in fixed T3 or T4 rectal cancer is the only exception as conversion of a primarily "unresectable" tumour to a R0 resectable one was the primary aim[32]. The PROSPECT phase 2/3 randomised trial (NCT01515787) compares standard preoperative CRT versus preoperative chemotherapy with selective CRT restricted to poor responders. This trial incorporated R0-resection and time to LRR as early stopping criteria for the phase 2 part, and proceeded to phase 3, when stopping criteria were not met after evaluation of the first 366 patients. Co-primary endpoints of the phase 3 part are time to LRR and DFS. Positive resection margins have been substantially reduced (<5-7%) with modern MRI staging, and improved CRT/SCRT and surgery, and may rather be used as benchmark quality measure than efficacy endpoint. The FOWARC phase 3 trial also assessed the value of preoperative chemotherapy alone as an alternative strategy in a 1:1:1 randomisation to either 5-FU-CRT followed by surgery and adjuvant 5-FU, or the same treatment plus oxaliplatin on day 1 of each cycle, or four to six cycles of mFOLFOX6 followed by surgery and adjuvant mFOLFOX6 [33]. Despite the significantly higher rates of pCR and tumour downstaging in the CRT arms, the primary endpoint, 3-year DFS was similar between the arms [33,34]. This study had important limitations including lack of data on the use of radiotherapy in T4 or CRM+ tumours, lack of a formal non-inferiority hypothesis, high rates of protocol violation or lost to follow-up (18%), and a dropout of 13%[33,34].

Sphincter-sparing surgery indirectly reflects tumour response to preoperative treatment and was used in the Polish I trial as the primary endpoint[35]. Despite better response in the CRT arm with higher pCR rates and less CRM involvement, no improvement in sphincter preservation was achieved. Note that eligibility for the Polish I trial was extended to patients with tumours of the mid rectum which could have diluted treatment effects. Moreover, the decision to conduct a sphincter-sparing surgery depends on tumour size, location, patient-related factors, surgical expertise and preferences.

Early endpoints used for limited- or non-operative management

Complete clinical response (cCR) has been introduced as a clinical endpoint following the implementation of organ preservation (**Figure 2**). Habr-Gama and colleagues were the first to pioneer the selective nonoperative management (NOM) approach in selected patients with cCR after conventional 5-FU CRT[36]. The International Watch and Wait Database reported the clinical outcome in 880 patients with cCR managed by NOM after neoadjuvant treatment in 47 institutions in 15 countries[37]. After a median follow-up time of 3.3 years, the 2-year cumulative incidence of local regrowth was 25.2%, of which 97% were intraluminal. The 5-year OS and DSS were 85% and 94%[37]. Similar promising data have been provided by the Manchester group [38], the Maastricht/NKI Amsterdam study[39] and the MSKCC database[40]. There is, however, currently no international consensus on how to best define cCR to facilitate consistency among current and future clinical trials. Imaging modalities, namely T2 and diffusion weighted MRI, PET-CT, and endoscopic ultrasound, have limitations to accurately predict pCR and/or to identify patients with cCR.

The Maastricht/NKI group previously provided a pragmatic definition for cCR as follows: 1) Substantial downsizing with no residual tumour or residual fibrosis only (with low signal on diffusion-weighted imaging MRI), sometimes associated with residual wall thickening due to edema; 2) no suspicious lymph nodes on MRI; 3) no residual tumour at endoscopy or only a small residual erythematous ulcer or scar; 4) negative biopsies from the scar, ulcer, or former tumour location (not mandatory); and 5) no palpable tumour, if initially palpable[39]. First assessment should be initiated 6-8 weeks after CRT and then performed 3-monthly. Importantly, many tumors are not reachable during digital rectal examination (DRE), and as opposed to endoscopic or MRI images, DRE cannot be objectively documented. In addition, endoscopy and MRI do not always correlate well with each other, which adds to the complexity and makes interpretation of response in the NOM setting challenging[39]. The varying definition of cCR and near-cCR, based on clinical, endoscopic and imaging criteria, and the follow-up protocol used after NOM are shown in **Supplementary Material and Supplementary Table 1-2.** Randomised trials currently testing the concept of organ preservation by NOM or local excision (LE)/transanal endoscopic microsurgery (TEM), and their corresponding endpoints are shown in **Table 2.**

Intermediate and long-term endpoints after surgery or limited-/non-operative management

In the past, **locoregional recurrence** (LRR) constituted the most important form of treatment failure in rectal cancer. Preoperative CRT/SCRT and improved TME quality have reduced LRR considerably. The early Swedish trial was performed before the routine use of TME and randomised patients to preoperative SCRT versus surgery alone. LRR decreased from 27% to 11% after addition of SCRT[41], which translated into a significant OS benefit. The Dutch Trial investigated preoperative SCRT plus TME vs TME alone, and the MRC CR07 trials compared preoperative SCRT against selected postoperative CRT in patients with positive CRM. Both trials used LRR as the primary endpoint and confirmed a significant advantage for SCRT. However, the numerically smaller differences in LRR (11% with surgery alone vs 4-5% in the preoperative SCRT groups at 3 years) did not translate into an OS benefit, whereas DFS was significantly improved in the MRC CR07 trial[42]. The Stockholm III trial confirmed non-inferiority of SCRT with immediate versus delayed surgery or long-course RT using time to LRR as the primary endpoint[43].

LRR as a primary endpoint in modern trials has been criticised because of the low incidence of events and the need for long term follow-up to identify late recurrences. The NSABP R-04 phase 3 trial compared preoperative CRT with capecitabine versus 5-FU, both with or without oxaliplatin. Three-year rates of LRR, the primary endpoint, among patients who underwent R0 resection ranged from 3·1 to 5·1%, with no significant differences between treatment groups[44].

Local regrowth occurs in approximately 20-30% of patients with initial cCR managed by NOM. Unlike LRR after radical surgery, which are typically extraluminal, difficult to salvage, and may occur beyond 5 years, local regrowth after cCR mostly occurs within 2 years of follow-up, is almost always intraluminal, and can be easily salvaged by curative surgery[37]. Thus, randomised trials testing NOM or LE/TEM use 12-months or 3-year organ preservation rates or DFS as primary endpoint (**Table 2, Figure 2**).

Distant control has not been used as a primary endpoint in phase 3 trials. Despite the improvement in LRR after preoperative (C)RT, OS was not improved in the CAO/ARO/AIO-94 and the Dutch TME trials[1,2]. New chemotherapy agents including oxaliplatin and irinotecan have been tested as part of phase 1-3 trials of preoperative CRT. A recent meta-analysis evaluated the addition of platinum derivates to 5-FU-based neoadjuvant CRT for rectal cancer. In 10 randomised controlled trials with 5599 patients, addition of oxaliplatin to CRT led to a significantly increased pCR (P=0·002) and reduced distant recurrence (P=0·004), however,

benefits were accompanied by higher rates of grade 3-4 toxicities without significant improvements in DFS (P=0·07) or OS (P=0·23)[45].

Disease-free survival (DFS) has been defined in adjuvant treatment trials for colon cancer as the time from randomisation to local or distant recurrence, second cancer, or death from any cause [46]. In a meta-analysis, the benefit of adjuvant treatments on 3-year DFS correlated significantly with the benefit on 5-year OS[47], and 3-year DFS has been accepted as a surrogate for OS in resectable colon cancer. Although 2-year DFS was a stronger predictor for OS than pCR after preoperative treatment among 2795 rectal cancer patients across 5 phase 3 studies[48], it has not been adopted as an endpoint. The definition of DFS has varied considerably among the different rectal cancer multimodal trials, especially with regard to surgery. Some groups included only patients receiving radical surgery (R0), whereas others also included patients with progressive disease before surgery, incomplete (R2) or no surgery in the definition. The role of adjuvant chemotherapy following standard neoadjuvant CRT/SCRT was assessed within 5 randomised trials (**Supplemetary Table 3**) but remains unclear as most of these studies suffered from low accrual, poor treatment compliance and suboptimal regimens, among others, and apart from the small ADORE trial, failed to demonstrate a clinical benefit for the primary endpoints, DFS or OS[49].

Composite endpoints

Composite endpoints have the advantage of reducing the required sample size and cost of a trial by increasing the event rates and should ideally incorporate components that occur with similar frequency, are of comparable importance to patients, and are affected to a similar degree by the intervention. Criticism has been raised that these conditions are unlikely to be met in practice[50,51]. The GRECCAR-2 phase 3 trial randomised 148 patients with T2/T3 lower rectal tumours and residual tumour ≤2 cm after CRT to receive either LE or TME[52]. The primary endpoint was the occurrence of any of the following untoward events: death, local or distant recurrence, severe surgical complications or major morbidity 2 years after surgery. Patients with ypT0-1 in the LE arm were followed-up, whereas a completion TME was required for ypT2-3 tumours. 35% of patients in the LE group needed a completion TME that significantly increased morbidity and side-effects. At 2 years, one or more events occurred in 56% and 48% of patients in the LE and TME arm, respectively (P=0·430). Similar 3-year LR rates were observed between the two arms (5% vs. 6%, P=0·680)[52]. The composite endpoint of this trial

was novel but its interpretation remains challenging due to competing risks between the different individual endpoint components[50,51].

Patient-reported outcome measures

PROMs were implemented to study the impact of the disease and/or treatment on **QoL**. Few phase 3 trials have reported QoL in rectal cancer. Marijnen et al. examined QoL in patients treated with TME with or without SCRT as part of the Dutch trial in 990 disease free patients up till two years after randomization. Few differences were found in QoL between patients treated with or without SCRT[53]. In a follow-up study, Peeters et al. examined toxicity using an in-house questionnaire in 597 patients[54]. After a median follow-up of 5.1 years, SCRT resulted in significantly higher rates of fecal incontinence with pad wearing and bowel dysfunction and worse sexual functioning[53,54]. The Nordic trial that compared QoL using the EORTC QLQ-C30 questionaire after CRT vs RT alone reported higher incidence of social dysfunctioning, dyspnea and diarrhea in the CRT arm but similar overall QoL in both groups[55]. In the Polish I trial that tested SCRT vs CRT, no significant differences in anorectal and sexual functions or QoL based on QLQ-C30 were observed between the two arms[56]. Similarly, the ACCORD 12/PRODIGE 2 phase 3 trial that randomised patients to capecitabine CRT with or without oxaliplatin showed similar QoL in both arms using the QLQ-C30 and -CR38 inventories [57].

DISCUSSION

The Definition for the Assessment of Time-to-event Endpoints in CANcer trials (DATECAN) project has been developed to provide consensus-based recommendations for clinical endpoints and facilitate consistency in reporting clinical data[58], but no consensus regarding the choice of endpoint in the different phase trials exists to date for rectal cancer. Historically, there has been a large variability regarding the choice of primary endpoints in phase 2 and phase 3 trials, leading to inter-trial inconsistency and difficulty in data interpretation. The advantages and disadvantages of the different clinical endpoints are summarized in **Supplementary Table 4**.

Undoubtedly, it is unlikely that we will ever be able to establish the perfect clinical endpoint for a trial, as all endpoints have advantages and disadvantages. Despite these limitations, it is important to define the most appropriate endpoint(s) for the respective trial phases to advance

progress. For rectal cancer, we propose a pragmatic approach to tailor measurement of efficacy to the specific clinical question to be addressed as summarized below:

- **Phase 1 trials:** Assessment of dose-limiting toxicities and definition of the recommended dose level for phase 2 testing remain standard primary endpoints. Unlike phase 1 trials performed in patients with metastatic disease refractory to prior lines of treatment, phase 1 trials in rectal cancer are commonly performed in treatment-naïve patients treated with curative intent. Thus, when testing the toxicity and feasibility of new chemotherapy, targeted or immunotherapy agents to standard CRT/SCRT, and/or RT dose escalation, more complex phase 1 designs such as the time-to-event continual reassessment method (TITE-CRM)[59] may be indicated. Monitoring of tumour response (especially exceptional responders) and molecular profiling to dissect response signals should be integral parts of phase 1 testing.
- **Phase 2 trials:** In phase 2 trials where CRT/SCRT are used preoperatively followed by radical surgery, NAR and TRG might be more appropriate than pCR for three reasons: (1) pCR did not correlate with improved OS in surrogate analyses, and was inferior to other endpoints, such as 2-year DFS[11,12,48]; (2) in contrast to pCR that represents a binary histological parameter, NAR and TRG might reflect treatment response and cancer biology better as they reflect a continuum of tumour regression; (3) NAR and TRG have been validated as individual-level surrogate markers for DFS[22,28]. The NAR score is attractive as it takes pretreatment cT-category into account, incorporates the prognostically most relevant ypN-categories, and had slightly better prognostic value for OS compared to pCR in rectal cancer trials[20,23]; NAR is not characterised by the inherent limitations of TRG assessment, including lack of a universally approved classification system.

  In phase 2 trials with the aim of NOM, cCR and sustained cCR during follow-up need to be reported, in conjuction with local regrowth rates and results of salvage surgery. If a LE/TEM is performed for near cCR further management is guided by histopathologic asessement, although criteria for completion TME versus follow-up (e.g., ypT0-1 vs ypT2-3) require further validation. For both NOM and LE/TEM organ preservation without non-salvageable locally progressive disease or permanent stoma is the most meaningful endpoint. Also, there is a need to develop a validated measure for anorectal function after NOM and LE/TEM.

  Single arm phase 2 trials are flawed by a necessary comparison with historical controls, whereas many factors can change – not least the quality of imaging, surgery and

pathology, and time intervals for response assessment. Also, quality control of imaging reads, especially pelvic MRI, is important to facilitate precise staging and response assessment. Randomised phase 2 trials will, to a certain extent, overcome these shortcomings. A "pick the winner", randomised phase 2 design was recently adopted by the CAO/ARO/AIO-12[18] and the OPRA trial for TNT (NCT02008656), the GEMCAD 1402[60] trial for inclusion of aflibercept, and the NRG GI002 trial (NCT02921256) for testing new radiosensitisers.

- **Phase 3 trials:** DFS constitutes the most suitable primary clinical endpoint, despite the lack of validation as surrogate for OS in rectal cancer. In the era of organ preservation, it is essential to incorporate the NOM into DFS for patients with cCR. **Table 4** summarizes the definition of events we propose to be appropriate to calculate DFS in multimodal rectal cancer phase 3 trials. This proposal is based on the consensus agreement for DFS in the adjuvant setting for colon cancer[46], but additionally includes events occurring during neoadjuvant treatment and at surgery, also incorporating the NOM and LE/TEM approach.

Future perspectives

The choice of primary endpoint in clinical trials remains challenging. Relying on surrogates and intermediate endpoints always carries the risk of conducting trials that might not improve long-term clinical outcome. The potential biological mechanisms that could underline poor correlation of surrogate endpoints with survival in some studies remain unclear. In the context of tumor heterogeneity, it is not unlikely that in some patients the local treatment effect on the primary tumour, as reflected by pCR, might not fully encapsulate the tumor cell propensity (or lack of) for micrometastatic seeding[61].

In the era of personalized medicine, the opinion of the patient needs to be considered and patients should be increasingly engaged in decision making. Future trials should aim to use endpoints, the relevance of which has preferably been agreed between patients, clinicians and regulatory authorities, as exemplified by the COMET (core outcome measures in effectiveness trial) and the CORMAC (core outcome set for clinical trials of CRT intervention) initiatives in anal cancer[62,63]. More data on PROMs are needed, especially in the increasingly-adopted setting of NOM. Randomisation should continue to constitute the reference method for phase 2 and phase 3 trials to provide robust evidence on the efficacy of new treatments.

Recently, modern biological assays such as measurement of circulating tumour and free DNA (ctDNA and cfDNA) have been explored in locally-advanced rectal cancer to identify patients at risk to develop metastases and tailor adjuvant chemotherapy[64,65]. Despite initial promising findings in prospective cohorts, these data need to be validated in further studies before considering implementation in the clinical setting.

Altogether, efforts should be made to continue evaluating surrogate and clinical endpoints using robust statistical methods. We recommend that data from large randomised clinical trials in rectal cancer become available to establish an international database, as suggested for other clinical settings[66] to allow access to all interested parties and stakeholders. Such a database will help provide important answers on the value of surrogate and clinical endpoints in the best interest of patients with rectal cancer.

Conflict of interest: The authors declared no conflicts of interest related to the present manuscript.

Author contributions: All authors (EF, RGJ, AA, RGT, GB, KH, CM, BDM, EL, PQ, DSM, JGA, MAG, VV, MB, CR) contributed to literature search, writing or review of the manuscript, and approved the final manuscript. EF, RGJ are joint first authors. MB and CR are joint last authors.

NIH Funding: Bruce Minsky has obtained funding from the NIH, US for research purposes that are not related to the present manuscript (Project numbers: 2U19CA021239-35 NCI, NIH/NCI; 1 U10 CA180858-01, NIH/NCI).

REFERENCES

1. van Gijn W, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. Lancet Oncol 2011; **12**(6): 575-82.
2. Sauer R, Liersch T, Merkel S, et al. Preoperative versus postoperative chemoradiotherapy for locally advanced rectal cancer: results of the German CAO/ARO/AIO-94 randomized phase III trial after a median follow-up of 11 years. J Clin Oncol 2012; **30**(16): 1926-33.
3. Wilson MK, Collyar D, Chingos DT, et al. Outcomes and endpoints in cancer trials: bridging the divide. Lancet Oncol 2015; **16**(1): e43-52.
4. Buyse M, Sargent DJ, Grothey A, Matheson A, de Gramont A. Biomarkers and surrogate end points--the challenge of statistical validation. Nat Rev Clin Oncol 2010; **7**(6): 309-17.
5. Bujko K, Nowacki MP, Nasierowska-Guttmejer A, Michalski W, Bebenek M, Kryj M. Long-term results of a randomized trial comparing preoperative short-course radiotherapy with preoperative conventionally fractionated chemoradiation for rectal cancer. Br J Surg 2006; **93**(10): 1215-23.
6. Ngan SY, Burmeister B, Fisher RJ, et al. Randomized trial of short-course radiotherapy versus long-course chemoradiation comparing rates of local recurrence in patients with T3 rectal cancer: Trans-Tasman Radiation Oncology Group trial 01.04. J Clin Oncol 2012; **30**(31): 3827-33.
7. Prentice RL. Surrogate endpoints in clinical trials: definition and operational criteria. Stat Med 1989; **8**(4): 431-40.
8. Buyse M, Molenberghs G, Paoletti X, et al. Statistical evaluation of surrogate endpoints with examples from cancer clinical trials. Biom J 2016; **58**(1): 104-32.
9. Fokas E, Liersch T, Fietkau R, et al. Downstage migration after neoadjuvant chemoradiotherapy for rectal cancer: the reverse of the Will Rogers phenomenon? Cancer 2015; **121**(11): 1724-7.
10. Maas M, Nelemans PJ, Valentini V, et al. Long-term outcome in patients with a pathological complete response after chemoradiation for rectal cancer: a pooled analysis of individual patient data. Lancet Oncol 2010; **11**(9): 835-44.
11. Bonnetain F, Bosset JF, Gerard JP, et al. What is the clinical benefit of preoperative chemoradiotherapy with 5FU/leucovorin for T3-4 rectal cancer in a pooled analysis of EORTC 22921 and FFCD 9203 trials: surrogacy in question? Eur J Cancer 2012; **48**(12): 1781-90.

12. Petrelli F, Borgonovo K, Cabiddu M, Ghilardi M, Lonati V, Barni S. Pathologic complete response and disease-free survival are not surrogate endpoints for 5-year survival in rectal cancer: an analysis of 22 randomized trials. J Gastrointest Oncol 2017; **8**(1): 39-48.

13. Francois Y, Nemoz CJ, Baulieux J, et al. Influence of the interval between preoperative radiation therapy and surgery on downstaging and on the rate of sphincter-sparing surgery for rectal cancer: the Lyon R90-01 randomized trial. J Clin Oncol 1999; **17**(8): 2396.

14. Lefevre JH, Mineur L, Kotti S, et al. Effect of Interval (7 or 11 weeks) Between Neoadjuvant Radiochemotherapy and Surgery on Complete Pathologic Response in Rectal Cancer: A Multicenter, Randomized, Controlled Trial (GRECCAR-6). J Clin Oncol 2016 34(31):3773-3780.

15. Evans J BJ, Sizer B, et al. Results of a prospective randomised control 6 vs 12 trial: Is greater tumour downstaging observed on post treatment MRI if surgery is delayed to 12-weeks versus 6-weeks after completion of neoadjuvant chemoradiotherapy? Ann Oncol 2016, 27: suppl_6, 452O.

16. Akgun E, Caliskan C, Bozbiyik O, et al. Randomized clinical trial of short or long interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer. Br J Surg 2018; **105**(11): 1417-25.

17. Garcia-Aguilar J, Chow OS, Smith DD, et al. Effect of adding mFOLFOX6 after neoadjuvant chemoradiation in locally advanced rectal cancer: a multicentre, phase 2 trial. Lancet Oncol 2015; **16**(8): 957-66.

18. Fokas E, Allgauer M, Polat B, et al. Randomized Phase II Trial of Chemoradiotherapy Plus Induction or Consolidation Chemotherapy as Total Neoadjuvant Therapy for Locally Advanced Rectal Cancer: CAO/ARO/AIO-12. J Clin Oncol 2019: JCO1900308.

19. Fokas E, Liersch T, Fietkau R, et al. Tumor regression grading after preoperative chemoradiotherapy for locally advanced rectal carcinoma revisited: updated results of the CAO/ARO/AIO-94 trial. J Clin Oncol 2014; **32**(15): 1554-62.

20. George TJ, Jr., Allegra CJ, Yothers G. Neoadjuvant Rectal (NAR) Score: a New Surrogate Endpoint in Rectal Cancer Clinical Trials. Curr Colorectal Cancer Rep 2015; **11**(5): 275-80.

21. Valentini V, van Stiphout RG, Lammering G, et al. Nomograms for predicting local recurrence, distant metastases, and overall survival for patients with locally advanced rectal cancer on the basis of European randomized clinical trials. J Clin Oncol 2011; **29**(23): 3163-72.

22. Fokas E, Fietkau R, Hartmann A, et al. Neoadjuvant rectal score as individual-level surrogate for disease-free survival in rectal cancer in the CAO/ARO/AIO-04 randomized phase III trial. Ann Oncol 2018; **29**(7): 1521-7.

23. Yothers G GT, Allegra CJ, et al. Predictive validity of NeoAdjuvant Rectal (NAR) Score and pathologic complete response (ypCR) for overall survival (OS) as surrogate endpoints in rectal cancer clinical trial. J Clin Oncol 2016; 34, no. 15_suppl: 3533-3.

24. Sclafani F, Kalaitzaki E, Cunningham D, Tait D, Brown G, Chau I. Neoadjuvant rectal score: run with the hare and hunt with the hounds. Ann Oncol 2018; **29**(11): 2261-2.

25. George TG YG, Hong TS, McGory Russell M, You YN, Parker W, Jacobs SA, Lucas PC, Gollub MJ, Hall WA, Kachnic LA, Vijayvergia N, Wolmark N. NRG-GI002: A phase II clinical trial platform using total neoadjuvant therapy (TNT) in locally advanced rectal cancer (LARC)—First experimental arm (EA) initial results. J Clin Oncol 2019; 37.15_suppl.3505.

26. van der Valk MJM, Vuijk FA, Putter H, van de Velde CJH, Beets GL, Hilling DE. Disqualification of Neoadjuvant Rectal Score Based on Data of 6596 Patients From the Netherlands Cancer Registry. Clin Colorectal Cancer 2019; **18**(2): e231-e6.

27. Trakarnsanga A, Gonen M, Shia J, et al. Comparison of tumor regression grade systems for locally advanced rectal cancer after multimodality treatment. J Natl Cancer Inst 2014; **106**(10).

28. Fokas E, Strobel P, Fietkau R, et al. Tumor Regression Grading After Preoperative Chemoradiotherapy as a Prognostic Factor and Individual-Level Surrogate for Disease-Free Survival in Rectal Cancer. J Natl Cancer Inst 2017; **109**(12).

29. Taylor FG, Quirke P, Heald RJ, et al. Preoperative magnetic resonance imaging assessment of circumferential resection margin predicts disease-free survival and local recurrence: 5-year follow-up results of the MERCURY study. J Clin Oncol 2014; **32**(1): 34-43.

30. Sclafani F, Brown G, Cunningham D, et al. Comparison between MRI and pathology in the assessment of tumour regression grade in rectal cancer. Brit J Cancer 2017; **117**(10): 1478-85.

31. Nagtegaal ID, Quirke P. What is the role for the circumferential margin in the modern treatment of rectal cancer? J Clin Oncol 2008; **26**(2): 303-12.

32. Bujko K, Wyrwicz L, Rutkowski A, et al. Long-course oxaliplatin-based preoperative chemoradiation versus 5 x 5 Gy and consolidation chemotherapy for cT4 or fixed cT3 rectal cancer: results of a randomized phase III study. Ann Oncol 2016; **27**(5): 834-42.

33. Deng Y, Chi P, Lan P, et al. Modified FOLFOX6 With or Without Radiation Versus Fluorouracil and Leucovorin With Radiation in Neoadjuvant Treatment of Locally Advanced

Rectal Cancer: Initial Results of the Chinese FOWARC Multicenter, Open-Label, Randomized Three-Arm Phase III Trial. J Clin Oncol 2016; **34**(27): 3300-7.

34. Deng Y, Chi P, Lan P, et al. Neoadjuvant Modified FOLFOX6 With or Without Radiation Versus Fluorouracil Plus Radiation for Locally Advanced Rectal Cancer: Final Results of the Chinese FOWARC Trial. J Clin Oncol 2019; **37**(34): 3223-33.

35. Bujko K, Nowacki MP, Nasierowska-Guttmejer A, et al. Sphincter preservation following preoperative radiotherapy for rectal cancer: report of a randomised trial comparing short-term radiotherapy vs. conventionally fractionated radiochemotherapy. Radiother Oncol 2004; **72**(1): 15-24.

36. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. Ann Surg 2004; **240**(4): 711-7; discussion 7-8.

37. van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. Lancet 2018; **391**(10139): 2537-45.

38. Renehan AG, Malcomson L, Emsley R, et al. Watch-and-wait approach versus surgical resection after chemoradiotherapy for patients with rectal cancer (the OnCoRe project): a propensity-score matched cohort analysis. Lancet Oncol 2016; **17**(2): 174-83.

39. Martens MH, Maas M, Heijnen LA, et al. Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. J Natl Cancer Inst 2016; **108**(12).

40. Smith JJ, Strombom P, Chow OS, et al. Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. JAMA Oncol 2019: e185896.

41. Swedish Rectal Cancer T, Cedermark B, Dahlberg M, et al. Improved survival with preoperative radiotherapy in resectable rectal cancer. N Engl J Med 1997; **336**(14): 980-7.

42. Sebag-Montefiore D, Stephens RJ, Steele R, et al. Preoperative radiotherapy versus selective postoperative chemoradiotherapy in patients with rectal cancer (MRC CR07 and NCIC-CTG C016): a multicentre, randomised trial. Lancet 2009; **373**(9666): 811-20.

43. Erlandsson J, Holm T, Pettersson D, et al. Optimal fractionation of preoperative radiotherapy and timing to surgery for rectal cancer (Stockholm III): a multicentre, randomised, non-blinded, phase 3, non-inferiority trial. Lancet Oncol 2017; **18**(3): 336-46.

44. Allegra CJ, Yothers G, O'Connell MJ, et al. Neoadjuvant 5-FU or Capecitabine Plus Radiation With or Without Oxaliplatin in Rectal Cancer Patients: A Phase III Randomized Clinical Trial. J Natl Cancer Inst 2015; **107**(11).

45. Huttner FJ, Probst P, Kalkum E, et al. Addition of platinum derivatives to fluoropyrimidine-based neoadjuvant chemoradiotherapy for stage II/III rectal cancer: systematic review and meta-analysis. J Natl Cancer Inst 2019; **111**(9):887-902.

46. Punt CJ, Buyse M, Kohne CH, et al. Endpoints in adjuvant treatment trials: a systematic review of the literature in colon cancer and proposed definitions for future trials. J Natl Cancer Inst 2007; **99**(13): 998-1003.

47. Sargent D, Shi Q, Yothers G, et al. Two or three year disease-free survival (DFS) as a primary end-point in stage III adjuvant colon cancer trials with fluoropyrimidines with or without oxaliplatin or irinotecan: data from 12,676 patients from MOSAIC, X-ACT, PETACC-3, C-06, C-07 and C89803. Eur J Cancer 2011; **47**(7): 990-6.

48. Valentini V, van Stiphout RG, Lammering G, et al. Selection of appropriate end-points (pCR vs 2yDFS) for tailoring treatments with prediction models in locally advanced rectal cancer. Radiother Oncol 2015; **114**(3): 302-9.

49. Carvalho C, Glynne-Jones R. Challenges behind proving efficacy of adjuvant chemotherapy after preoperative chemoradiation for rectal cancer. Lancet Oncol 2017; **18**(6): e354-e63.

50. Tomlinson G, Detsky AS. Composite end points in randomized trials: there is no free lunch. JAMA 2010; **303**(3): 267-8.

51. Buyse M. Generalized pairwise comparisons of prioritized outcomes in the two-sample problem. Stat Med 2010; **29**(30): 3245-57.

52. Rullier E, Rouanet P, Tuech JJ, et al. Organ preservation for rectal cancer (GRECCAR 2): a prospective, randomised, open-label, multicentre, phase 3 trial. Lancet 2017; 391(10116):122-123.

53. Marijnen CA, van de Velde CJ, Putter H, et al. Impact of short-term preoperative radiotherapy on health-related quality of life and sexual functioning in primary rectal cancer: report of a multicenter randomized trial. J Clin Oncol 2005; **23**(9): 1847-58.

54. Peeters KC, van de Velde CJ, Leer JW, et al. Late side effects of short-course preoperative radiotherapy combined with total mesorectal excision for rectal cancer: increased bowel dysfunction in irradiated patients--a Dutch colorectal cancer group study. J Clin Oncol 2005; **23**(25): 6199-206.

55. Braendengen M, Tveit KM, Hjermstad MJ, et al. Health-related quality of life (HRQoL) after multimodal treatment for primarily non-resectable rectal cancer. Long-term results from a phase III study. Eur J Cancer 2012; **48**(6): 813-9.

56. Pietrzak L, Bujko K, Nowacki MP, et al. Quality of life, anorectal and sexual functions after preoperative radiotherapy for rectal cancer: report of a randomised trial. Radiother Oncol 2007; **84**(3): 217-25.

57. Etienne PL FE, Gourgou S, et al. PRODIGE 2 phase III trial neoadjuvant in rectal cancer: Quality of life and results at 5 years. J Clin Oncol; 34, 15_suppl 3619-3619.

58. Bellera CA, Pulido M, Gourgou S, et al. Protocol of the Definition for the Assessment of Time-to-event Endpoints in CANcer trials (DATECAN) project: formal consensus method for the development of guidelines for standardised time-to-event endpoints' definitions in cancer clinical trials. Eur J Cancer 2013; **49**(4): 769-81.

59. de Haan R, van Werkhoven E, van den Heuvel MM, et al. Study protocols of three parallel phase 1 trials combining radical radiotherapy with the PARP inhibitor olaparib. BMC Cancer 2019; **19**(1): 901.

60. Fernandez-Martos C, Pericay C, Losa F, et al. Effect of Aflibercept Plus Modified FOLFOX6 Induction Chemotherapy Before Standard Chemoradiotherapy and Surgery in Patients With High-Risk Rectal Adenocarcinoma: The GEMCAD 1402 Randomized Clinical Trial. JAMA Oncol 2019 doi: 10.1001/jamaoncol.2019.2294.

61. Rose BS, Winer EP, Mamon HJ. Perils of the Pathologic Complete Response. J Clin Oncol 2016; **34**(33): 3959-62.

62. Glynne-Jones R, Adams R, Lopes A, Meadows H. Clinical endpoints in trials of chemoradiation for patients with anal cancer. Lancet Oncol 2017; **18**(4): e218-e27.

63. Fish R, Sanders C, Adams R, et al. A core outcome set for clinical trials of chemoradiotherapy interventions for anal cancer (CORMAC): a patient and health-care professional consensus. Lancet Gastroenterol Hepatol 2018; **3**(12): 865-73.

64. Tie J, Cohen JD, Wang Y, et al. Serial circulating tumour DNA analysis during multimodality treatment of locally advanced rectal cancer: a prospective biomarker study. Gut 2019; **68**(4): 663-71.

65. Khakoo S, Carter PD, Brown G, et al. MRI Tumor Regression Grade and Circulating Tumor DNA as Complementary Tools to Assess Response and Guide Therapy Adaptation in Rectal Cancer. Clin Cancer Res 2019.

66. Bertagnolli MM, Sartor O, Chabner BA, et al. Advantages of a Truly Open-Access Data-Sharing Model. N Engl J Med 2017; **376**(12): 1178-81.

67. Cotte E, Passot G, Decullier E, et al. Pathologic Response, When Increased by Longer Interval, Is a Marker but Not the Cause of Good Prognosis in Rectal Cancer: 17-year Follow-up of the Lyon R90-01 Randomized Trial. Int J Radiat Oncol Biol Phys 2016; **94**(3): 544-53.
68. Fernandez-Martos C, Pericay C, Aparicio J, et al. Phase II, randomized study of concomitant chemoradiotherapy followed by surgery and adjuvant capecitabine plus oxaliplatin (CAPOX) compared with induction CAPOX followed by concomitant chemoradiotherapy and surgery in magnetic resonance imaging-defined, locally advanced rectal cancer: Grupo cancer de recto 3 study. J Clin Oncol 2010; **28**(5): 859-65.
69. Dewdney A, Cunningham D, Tabernero J, et al. Multicenter randomized phase II clinical trial comparing neoadjuvant oxaliplatin, capecitabine, and preoperative radiotherapy with or without cetuximab followed by total mesorectal excision in patients with high-risk rectal cancer (EXPERT-C). J Clin Oncol 2012; **30**(14): 1620-7.
70. Rouanet P, Rullier E, Lelong B, et al. Tailored Treatment Strategy for Locally Advanced Rectal Carcinoma Based on the Tumor Response to Induction Chemotherapy: Preliminary Results of the French Phase II Multicenter GRECCAR4 Trial. Dis Colon Rectum 2017; **60**(7): 653-63.
71. Roh MS, Colangelo LH, O'Connell MJ, et al. Preoperative multimodality therapy improves disease-free survival in patients with carcinoma of the rectum: NSABP R-03. J Clin Oncol 2009; **27**(31): 5124-30.
72. Frykholm GJ, Pahlman L, Glimelius B. Combined chemo- and radiotherapy vs. radiotherapy alone in the treatment of primary, nonresectable adenocarcinoma of the rectum. Int J Radiat Oncol Biol Phys 2001; **50**(2): 427-34.
73. Kapiteijn E, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer. N Engl J Med 2001; **345**(9): 638-46.
74. Sauer R, Becker H, Hohenberger W, et al. Preoperative versus postoperative chemoradiotherapy for rectal cancer. N Engl J Med 2004; **351**(17): 1731-40.
75. Braendengen M, Tveit KM, Berglund A, et al. Randomized phase III study comparing preoperative radiotherapy with chemoradiotherapy in nonresectable rectal cancer. J Clin Oncol 2008; **26**(22): 3687-94.
76. Bosset JF, Calais G, Mineur L, et al. Fluorouracil-based adjuvant chemotherapy after preoperative chemoradiotherapy in rectal cancer: long-term results of the EORTC 22921 randomised study. Lancet Oncol 2014; **15**(2): 184-90.

77. Gerard JP, Conroy T, Bonnetain F, et al. Preoperative radiotherapy with or without concurrent fluorouracil and leucovorin in T3-4 rectal cancers: results of FFCD 9203. J Clin Oncol 2006; **24**(28): 4620-5.

78. Aschele C, Cionini L, Lonardi S, et al. Primary tumor response to preoperative chemoradiation with or without oxaliplatin in locally advanced rectal cancer: pathologic results of the STAR-01 randomized phase III trial. J Clin Oncol 2011; **29**(20): 2773-80.

79. Gerard JP, Azria D, Gourgou-Bourgade S, et al. Clinical outcome of the ACCORD 12/0405 PRODIGE 2 randomized trial in rectal cancer. J Clin Oncol 2012; **30**(36): 4558-65.

80. Schmoll HJ HK, Rice TJ, et al. Preoperative chemoradiotherapy and postoperative chemotherapy with capecitabine and oxaliplatin versus capecitabine alone in locally advanced rectal cancer: disease-free survival at interim analysis. Proc Am Soc Clin Oncol 2014;32: abstract 3501.

81. Hofheinz RD, Wenz F, Post S, et al. Chemoradiotherapy with capecitabine versus fluorouracil for locally advanced rectal cancer: a randomised, multicentre, non-inferiority, phase 3 trial. Lancet Oncol 2012; **13**(6): 579-88.

82. Rodel C, Graeven U, Fietkau R, et al. Oxaliplatin added to fluorouracil-based preoperative chemoradiotherapy and postoperative chemotherapy of locally advanced rectal cancer (the German CAO/ARO/AIO-04 study): final results of the multicentre, open-label, randomised, phase 3 trial. Lancet Oncol 2015; **16**(8): 979-89.

FIGURE LEGENDS

Figure 1. Illustration of the different outcome measures used in randomised clinical trials in patients with rectal cancer that received neoadjuvant treatment followed by TME surgery. The outcome measures used in randomised phase 3 clinical trials as primary endpoints are marked with orange colour. Non-coloured efficacy endpoints have been used as secondary outcome measures. The different preoperative treatment options that are characterised by variable length and time to surgery appear below the x-axis. Also, the timepoints of assessment of PROM/QoL as suggested in modern trials are shown here. Abbreviations: Txt, treatment; TME, total mesorectal excision; cTNM, clinical tumour/node/metastasis staging, MRF+/-, mesorectal fascia involvement, EMVI, extramural venous invasion; SI, sphincter involvement; PROM, patient reported outcome measures; QoL, quality of life; R0, complete resection; CRM+/-, circumferential resection margin involvement; TRG, tumour regression grading; pCR, pathological complete response; NAR, neoadjuvant rectal score; LRR, locoregional recurrence; DFS, disease-free survival; TdrTF, time to disease-related treatment failure; OS, overall survival; SCRT, short-course radiotherapy; CRT, chemoradiotherapy; TNT, total neoadjuvant therapy; S, surgery;

Figure 2. Illustration of the different outcome measures used in clinical trials in patients with rectal cancer that received neoadjuvant/definitive treatment followed by non-operative management or local excision. The different preoperative treatment options that are characterised by variable length and time to surgery appear below the x-axis. Of note, only the endpoints that have been used in randomised phase 2/3 clinical trials are illustrated; those used as primary endpoint are marked with orange color. The definition of organ preservation as endpoint is provided in **Table 2**. The composite endpoint refers to novel primary endpoint of the GRECCAR2 trial (see text, **Table 3**). Also, the timepoints of assessment of PROM/QoL as suggested in modern trials are shown here. Abbreviations: Txt, treatment; cTNM, clinical tumour/node/metastasis staging, MRF+/-, mesorectal fascia involvement, EMVI, extramural venous invasion; SI, sphincter involvement; NOM, non-operative management; LE, local excision; TME, total mesorectal excision; PROM, patient reported outcome measures; QoL, quality of life; MRF, mesorectal fascia; cCR, complete clinical response; mrTRG, magnetic resonance imaging-based tumour regression grading; R0, complete resection; DFS, disease-free survival; LRR, locoregional recurrence; SCRT, short-course radiotherapy; CRT, chemoradiotherapy; EBRT, external beam radiotherapy; Brachy, brachytherapy; * The follow-up protocol of NOM and definitions of cCR and near cCR as proposed by the Maastricht/NKI

group and others are shown in the main text, the **Supplementary material** and **Supplementary Tables 1-2**.