International consensus recommendations on key outcome measures of organ preservation in rectal cancer

Emmanouil Fokas[1, 2, 3, #], Ane Appelt[4, #], Robert Glynne-Jones[5], Geerard Beets[6,7], Rodrigo Perez[8], Julio Garcia-Aguilar[9], Eric Rullier[10], J. Joshua Smith[9], Corrie Marijnen[11], Femke P. Peters[11], Maxime van der Valk[6], Regina Beets-Tan[6,12], Arthur Sun Myint[13], Jean-Pierre Gerard[14], Simon P Bach[15], Michael Ghadimi[16], Ralf-Dieter Hofheinz[17], Krzysztof Bujko[18], Cihan Gani[19,20], Karin Haustermans[21], Bruce D. Minsky[22], Ethan Ludmir[22], Nick Paul West[23], Maria Antonietta Gambacorta[24], Vincenzo Valentini[24], Marc Buyse[25,26], Andrew G. Renehan[27], Alexandra Gilbert[4, §], David Sebag-Montefiore[4, §], Claus Rödel[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]Leeds Institute of Medical Research at St James's, University of Leeds, Leeds, UK

[5]Department of Radiotherapy, Mount Vernon Centre for Cancer Treatment, Northwood, Middlesex, UK

[6]GROW School for Oncology and Developmental Biology, Maastricht University, The Netherlands

[7]Department of Surgery, Netherlands Cancer Institute Amsterdam, Amsterdam, The Netherlands

[8]Department of Surgery, Angelita & Joaquim Institute, São Paulo, Brazil

[9]Colorectal Service, Department of Surgery, Memorial Sloan Kettering Cancer Center, New York, New York

[10]Department of Colorectal Surgery, Haut-Lévèque Hospital, CHU Bordeaux, France

[11]Department of Radiation Oncology, Netherlands Cancer Institute Amsterdam, Amsterdam, The Netherlands

[12]Department of Radiology, Netherlands Cancer Institute, Amsterdam, The Netherlands

[13]The Clatterbridge Cancer Centre, Liverpool, UK

[14]Service de Radiothérapie, Centre Antoine-Lacassagne, Nice, France

[15]Academic Department of Surgery, University of Birmingham, Birmingham, UK

[16]Department of General, Visceral, and Pediatric Surgery, University Medical Center, Göttingen, Germany

[17]Department of Medical Oncology, University Hospital Mannheim, University Heidelberg, Heidelberg, Germany

[18]Department of Radiotherapy, Maria Sklodowska-Curie National Research Institute of Oncology, Warsaw, Poland

[19]Department of Radiation Oncology, University Hospital and Medical Faculty Tübingen, Eberhard Karls University Tübingen, Germany
[20]German Cancer Research Center (DKFZ) Heidelberg and German Consortium for Translational Cancer Research (DKTK), Partner Site Tübingen, Tübingen, Germany
[21]Department of Radiation Oncology, University Hospital Leuven, Leuven, Belgium
[22]Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
[23]Division of Pathology and Data Analytics, Leeds Institute of Medical Research at St. James's, School of Medicine, Leeds University, Leeds, UK
[24]Department Radiation Oncology and Medical Oncology, Fondazione Policlinico Universitario A. Gemelli IRCCS, Università Cattolica del Sacro Cuore, Roma, Italy
[25]Interuniversity Institute for Biostatistics and Statistical Bioinformatics, Hasselt University, Diepenbeek, Belgium
[26]International Drug Development Institute, San Francisco, CA, USA
[27]Division of Cancer Sciences, School of Medical Sciences, Faculty of Biology, Medicine and Health, Manchester Academic Health Science Centre, University of Manchester, Manchester, UK; Colorectal and Peritoneal Oncology Centre, Christie NHS Foundation Trust, Manchester, UK

#joint first authors; §joint last authors

Type of article: Expert consensus statement

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

ABSTRACT

Multimodal treatment strategies for rectal cancer are increasingly embracing organ preservation, i.e. non-operative management or local excision, for patients with (near) clinical complete response after (chemo-)radiotherapy due to its oncological safety and reduction in surgical morbidity. However, standardisation of key outcome measures of organ preservation is lacking; this includes definition and choice of primary endpoints according to the trial phase and design, timepoint of response assessment, response-based decision, follow-up schedules, specific anorectal function tests, quality of life and patient reported outcomes. Thus, a consensus statement on outcome measures is necessary to ensure consistency and facilitate comparison between ongoing and future trials. Here, we have convened an international group of clinical trialists with extensive experience in rectal cancer management, including organ preservation, and used a Delphi process to establish the first international consensus recommendations of key outcome measures of organ preservation, to standardise reporting for trials and routine practice of organ preservation.

INTRODUCTION

There has been a progressive increase in the number of clinical trials examining organ preservation strategies, i.e. non-operative management (NOM) or local excision only (LE), after (chemo-)radiotherapy (CRT), in rectal cancer[1]. Habr-Gama and colleagues were the first to implement the selective NOM approach in patients with resectable rectal cancer who achieved a clinical complete response (cCR) following CRT[2]. Since then, several studies, including the international Watch and Wait database (IWWD) analysis, have shown that deferral of surgery in patients with cCR appears to be oncologically safe; although more randomised data are needed to confirm long-term oncological safety and superiority of organ preservation regarding quality of life (QoL) assessed by patient reported outcomes (PROs)[3-11]. Local excision (LE) by transanal endoscopic microsurgery (TEM) or transanal minimally invasive surgery (TAMIS) is an alternative organ preservation strategy approach for selected patients with small T1-T3 low rectal cancer and good response after CRT, as shown in CARTS, TREC and GRECCAR2 trials[9,12-14]. The STAR-TREC trial (NCT02945566) is exploring NOM and LE, depending on the degree of response after neoadjuvant treatment in early stage disease.

Reflective of the Definition for the Assessment of Time-to-event Endpoints in CANcer trials (DATECAN) project[15], we recently provided recommendations for the use of clinical and surrogate endpoints in the different phases (1-3) of rectal cancer trials[16]. However, standardization of key outcome measures of organ preservation is lacking in rectal cancer. Organ preservation trials are characterised by marked heterogeneity in selection criteria, treatment strategies, choice of endpoint and design that limit data interpretation and comparison between studies. Hence, an international consensus is needed to ensure consistency, and facilitate appropriate data collection, interpretation and outcome comparison for organ preservation as part of trials (i.e. "intended" organ preservation) and outside trials (i.e. "incidental" organ preservation in patients with cCR after standard treatment that is now permitted by several guidelines including ESMO,[17] NCCN[18], ASTRO[19]). Here, we aim to establish the first clinical expert consensus statement on key outcome measures for organ preservation in rectal cancer, with a particular focus on NOM. We have convened an international group of clinical trialists with extensive experience in rectal cancer studies, including organ preservation strategies, and used the Delphi process to collect opinions, with the aim to standardise measurement and reporting in this setting.

METHODS

Literature search strategy and selection criteria

References were retrieved from four electronic databases (PubMed, MEDLINE, Web of Science, and the Cochrane Library, Google Scholar) for published articles and abstracts from international meetings from retrospective, prospective and randomised clinical trials investigating organ preservation for rectal cancer, published from inception to 1 April 2020. The literature search criteria and method are described in detail in **Supplementary Methods**. Two investigators (EF and CR) extracted the key outcome measures of organ preservation from all selected studies to be included into the Delphi process for consensus statement and standardisation, reviewed the list of retrieved articles and selected potentially relevant articles. The flowchart of article selection process is shown in **Figure 1**.

Formation of consensus panel and Delphi method to establish a consensus

The guideline panel comprised a multidisciplinary and interprofessional team, including clinical oncologists, radiation oncologists, medical oncologists, surgical oncologists, pathologist, radiologists with expertise in rectal cancer as well as bioinformatician. A Delphi method was used to vote to achieve consensus statements by all panelists using the SurveyMonkey program (https://www.surveymonkey.com) and electronic communications. To reach a consensus on the different outcome measures, a threshold of 70% or more for agreement was required for each item. The formation of consensus panel and Delphi method are described in detail in **Supplementary Methods.**

RESULTS

Literature search and review

The literature search retrieved 3090 publications. 667 abstracts were selected for full-text assessment, after removal of duplicates, and screening of the title and abstract (**Figure 1**). After full-text article review and exclusion of manuscripts unrelated to the present topic and non-English articles, 396 manuscripts were considered relevant to the scope of the present study. We identified the following 7 outcome measures as key to an organ preservation strategy: definition of endpoints (methodology and criteria to define response, unequivocal nomenclature); choice of primary endpoint according to the trial phase and design; timepoint of tumour response assessment (RA) to determine cCR; response-based decision algorithms and use of biopsy; follow-up methods (schedules and timelines); organ preservation-specific

anorectal function test; QoL assessment and PROs. The 7 outcome measures were then developed into 32 clinical questions to include in the Delphi survey (**Supplementary Table 1**).

Consensus procedure and Delphi rounds

The questionnaires of 1st and 2nd Delphi round as (R1 and R2) on the 7 key outcome measures of organ preservation together with the corresponding answers are provided in **Supplementary Tables 1-2**, respectively. In the 3rd round (R3), the final consensus manuscript recommendations for the key outcome measures were prepared and agreed upon by all members (100%) of the panel. The flow diagram of the study procedures including R1 to R3 to establish an international consensus is shown in **Figure 2**. The results of the consensus procedure and Delphi rounds are described in detail in **Supplementary Results**.

CONSENSUS STATEMENT RECOMMENDATIONS

Criteria, definition and nomenclature of clinical endpoints

Table 1 summarises the definitions of the different clinical endpoints after consensus recommendation was achieved. The panel reached a consensus as part of the Delphi process and agreed upon the definitions of organ preservation, locoregional regrowth after NOM and locoregional recurrence after LE or total mesorectal excision (TME), respectively. Definitions of incomplete/poor response, local regrowth and local recurrence were provided separately for clarity. The various criteria reported in the literature to define cCR are shown in **Supplementary Table 4.** The panel recommended that the "Amsterdam/Maastricht" criteria[4] were best suited to define cCR and near cCR (ncCR). The panel also agreed with the definition of organ-preservation-adapted DFS, as proposed recently[16]. The definition of TME-free DFS used in the OPRA trial was introduced for the first time in the literature at ASCO 2020[20,21], which explains why consensus was not reached for this endpoint. As such, the definition of TME-free DFS was provided separately by the primary investigator of the OPRA trial (JGA).

Choice of primary endpoint according to the trial phase and design

The panel recommended that different primary endpoints should be used according to the trial design, taking into consideration the initial tumour stage, use of standard or intensified experimental treatment regimen, intended or incidental organ preservation, NOM or LE strategies, and overall aim. The primary endpoints that reached consensus after the Delphi process according to the different trial designs together with representative trial examples are described below:

- **Early tumour response assessment (i.e. cCR rate)** should be used as primary endpoint for early phase 1/2 trials intentionally aiming to increase cCR rates and enable NOM/LE by more intense RT/CRT/total neoadjuvant treatment (TNT) regimens; to select tolerable and locally effective treatment regimens for further testing in larger scale trials (e.g. Danish trial[7], CAO/ARO/AIO-16 trial: NCT03561142). The risks and benefits of more intense treatments should be considered carefully.
- **Organ preservation assessed at 30-36 months after treatment start** as an intermediate endpoint should be the primary endpoint for (randomised) phase 2/3 trials using either NOM or LE (for cCR or ncCR) (e.g. WW3, STAR-TREC, ACO/ARO/AIO-18.1 trials). Function, toxicity and QoL were regarded as pivotal secondary outcomes, to be considered for inclusion as composite or co-primary endpoints (e.g. GRECCAR2 trial[9,12]).
- **Organ preservation-adapted DFS at 3 years**[16] should be used as a primary endpoint if organ preservation is allowed within but is not the primary purpose of a (late) phase 3 trial, especially in locally-advanced tumours.

Relevant to this recommendation, **Table 2** only includes randomised studies of organ preservation showing both the variability among studies regarding the timepoint of response assessment (RA) to determine cCR (discussed below) as well as the primary endpoint selected.

Timepoint of early tumour RA to determine cCR

- Although the evidence on optimal timing for RA to determine cCR is still growing and influenced by many variables (such as initial tumour stage, biology, treatment duration and intensity, interval from treatment completion, methodology to assess response etc.), the panel indicated the importance providing clear recommendations for future trials and routine practice that achieved consensus. The panel consensus recommendation on the timepoint of RA and determining cCR according to treatment design is summarised in **Table 3**. Representative trial examples illustrating the complexity of accurate timing for assessing response due to the highly variable treatment design and duration among the different clinical trials are shown in **Figure 3** (and **Table 2** that only shows randomised studies).

Response-based decision and use of biopsy

A question commonly raised is whether clinicians should wait longer before deciding on surgery if restaging after preoperative treatment shows ncCR. While timing for evaluation of

cCR greatly depends on the context of treatment design, the panel supported longer waiting in this setting, although no consensus was reached on the timing of the second assessment. Notably, this decision should be made also considering initial stage, trial treatment design and duration for RA, as described above.

Another important point concerned the role of biopsy in case of ncCR or cCR. In both cases, there was consensus agreement that biopsy does not provide additional value and could lead to false-negative results. Martens et al. followed-up the work from Maas et al. that clearly indicated the lack of added diagnostic value for biopsy[5,22]. Thus, a biopsy is not mandatory or recommended by the panel. In the case where a biopsy is nevertheless performed in a patient with ncCR and is negative, the panel recommended that longer waiting and reassessment after 6-12 weeks could be considered, again depending on the treatment design.

Follow-up procedures and schedule

The panel reached a consensus that CEA, digital rectal examination (DRE), rectoscopy, pelvic MRI and chest/abdomen CT should be part of the follow-up for organ preservation. The majority indicated that CEA should be assessed every 3 months during years 1-3, and every 6 months at years 4-5 after completion of treatment for organ preservation. Consensus was established that DRE, endoscopy and MRI should be conducted every 3-4 months during years 1-2, and every 6 months in years 3-5. Finally, the preferred time schedule to perform CT-thorax/abdomen is every 6-12 months at year 1, and every 12 months during years 2-5. The follow-up procedures and schedule that reached consensus is shown in **Table 4**.

Anorectal function measurement

The panel was asked to select among commonly used tests to measure anorectal function, combining a mix of clinician and patient reported instruments. These included the Wexner score[23], the Low Anterior Resection Syndrome (LARS) score[24], the MSKCC Bowel Function Instrument (MSKCC BFI) score[25], the Vaizey score[26] and manometry (**Supplementary Table 5**). The LARS score (PRO) received most votes and reached consensus. Participants indicated that, together with available methods, a new score specific to the organ preservation should be developed; commenting on the need to measure urinary and sexual dysfunction in addition to bowel dysfunction.

QoL assessment and PROs

The panel achieved a consensus that EORTC QLQ-C30 should always be used. The panel was asked to vote on 5 proposed QoL and function scales to be recorded. These included overall QoL, physical function, role function, social function and emotional function. Consensus was achieved for all 5 proposed QoL and function scales.

The panel also agreed on the 10 most important symptomatic toxicity items among a list of 20 proposed items for evaluation as part of a patient-reported assessment. These included bowel urgency, fecal incontinence, bowel frequency, diarrhea, tenesmus, toilet dependency, night time bowel opening, urinary urgency, impotence and pain. 42% voted for the use of the EORTC QLQ-CR29 in addition to QLQ-C30. The EORTC-QLQ CR29 although covers many bowel, urinary, stoma and sexual issues, does not include all bowel symptoms experienced following NOM/LE, in particular bowel urgency and toilet dependency. Although these bowel issues are included in the LARS score, it lacks items on urinary and sexual dysfunction, and stoma-related items for patients who fail to achieve organ preservation. All participants indicated the need for developing a new, validated PRO (or extension) specific for NOM/LE (**Supplementary Table 1**).

Finally, the panel was provided a list with different timepoints to vote the optimal timings for measurement of symptomatic toxicity, QoL and function. The panel recommended that toxicity should be measured at baseline, 3 months, 12 months, 24 months, 36 months and 60 months after decision for NOM/LE. A similar consensus was reached by the panel for the same timepoints for QOL and function measurement.

DISCUSSION AND FUTURE PERSPECTIVES

We here provide the first international consensus recommendation on key outcome measures for organ preservation strategies in rectal cancer. Undoubtedly, we are still at a transitional phase, if not only the beginning of a new era, where evidence regarding many aspects of organ preservation is far from complete[1]. This is reflected by the inconsistency in reporting in clinical trials and retrospective or population-based series, which underlines the importance of the present study. Also, ambiguous clinical outcomes have often been reported, also due to heterogeneity in patient inclusion criteria for radiotherapy treatment and method, as well as chemotherapy regimen. We recommend that investigators use the consensus recommendation set as a framework for organ preservation in rectal cancer.

Ambiguous language in events defining clinical endpoints, such as cCR, regrowth, recurrence, organ preservation and DFS with or without considering regrowth has often led to confusion.

In the Champalimaud meeting it was agreed that the term "local regrowth" should replace local recurrence when tumour regrowth occurs after initial cCR, due to its different time course, salvageability and favourable prognosis[27]. Nevertheless, distinction between locoregional and local/regional regrowth (or recurrence) has been far from clear, and rigorous definitions were not provided. Here, exact description of endpoints reached consensus to avoid disparity, and enable future cross-trial comparisons. The recently-proposed improved definition of DFS (organ preservation-adapted DFS)[16] that incorporates NOM/LE reached consensus. Although TME-free DFS was only recently introduced as endpoint, reported in the OPRA trial at ASCO 2020[20,21], its definition was provided for future reference.

The choice of the most appropriate outcome measure is a crucial component of organ preservation trials[28]. Selection of primary endpoints in prospective studies has been rather arbitrary. Due to the different treatment strategies and duration, the panel acknowledged that "one size does not fit all" for organ preservation strategies, and recommended specific endpoints according to the clinical scenario. Similarly to the pCR endpoint in trials with radical surgery after neoadjuvant treatment[29], cCR was suggested as endpoint for early phase 1/2 trials using more intense RT/CRT/TNT regimen to select tolerable and locally effective treatment regimens for further testing in larger scale trials (e.g. Appelt et al. in the Danish trial that used CRT followed by radiotherapy dose escalation with brachytherapy[7]). Of note, sustained cCR at 12 months is a part of the endpoint of organ preservation and was, thus, not recommended as a separate endpoint in the present consensus study. Instead, we proposed cCR as an early endpoint in small trials exploring promising regimens to achieve organ preservation, and not as the ultimate clinical endpoint. Organ preservation at 30-36 months after the start of treatment was agreed upon as the primary endpoint for phase 2/3 trials using NOM/LE to achieve organ preservation (as currently used in STAR-TREC, OPERA and ACO/ARO/AIO-18.1 trials). While the timepoint for defining organ preservation varies among studies (**Table 2**), a 30-36 month time window was recommended, reflecting the prolonged treatment time of TNT and that tumour regrowth mostly occurs up to 24-30 months after treatment completion[8,30]. Organ-preservation-adapted DFS was selected for phase 3 trials that allow organ preservation but specifically aim to improve oncological outcome, especially distant metastases (as in TRIGGER trial[31]).

There are no perfect primary endpoints for organ preservation as all endpoints are susceptible to pitfalls[32]. Also, the choice of primary endpoint serves the purpose of statistical trial design, whereas secondary endpoints, especially QoL and PROs (one of the main arguments for deferring surgery), should be regarded as equally important[13,33-35]. Shared decision making with

patients and risk-benefit analysis (e.g. balance between NOM/LE and treatment toxicity) should be considered for "intended" organ preservation trials. The fact that bad responders may receive overtreatment should not be underestimated, as shown in GRECCAR2 trial, where many patients in the LE group required completion TME, increasing morbidity and side-effects[9,12].

In that context, future studies should aim to clarify which inclusion criteria should be used to advocate LE, the optimal timing of LE depending on tumor response (cCR vs near cCR vs residual disease), and how this relates to pre-treatment staging[36-38].

The timepoint of determining cCR constitutes one of the biggest challenges, as tumour response to treatment is a dynamic phenomenon affected by tumour size, histology, biology, treatment strategy, and the time interval between preoperative/definitive treatment and decision for NOM/LE (or TME surgery)[16]. This is reflected in the variation of timepoint for RA among different studies due to the variation in treatment schedule and design (**Figure 3**). Knowledge on the kinetics of tumour response has mainly been derived from the operative setting. In a pooled analysis of 4431 patients, pCR rates increased with intervals greater than 6-7 weeks post-CRT, whereas the Dutch Surgical Colorectal Audit showed a peak in pCR at 10 weeks post-CRT i.e. 16 weeks after treatment start in 1593 patients[39]. The advent of TNT, with highly variable treatment duration among different trials, has added to the complexity of this issue. In a phase 2 trial, patients received two, four, or six cycles of FOLFOX chemotherapy after CRT, and underwent surgery at 6, 11, 15, and 19 weeks after completion of CRT; pCR rates were 18%, 25%, 30%, and 38%, respectively[40]. Whether these differences can be explained by the intensified chemotherapy or by the prolonged interval remains uncertain. The CAO/ARO/AIO-12 trial compared the two TNT sequences, induction CT/CRT vs CRT/consolidation CT, and demonstrated a pCR in 17% and 25%, respectively[41]. Similar data favouring the sequence CRT/CT were reported in the OPRA trial that showed 3-year TME-free survival rates of 59% vs 43% for CT/CRT[20].

The panel agreed that defining one specific time point for assessing cCR is impossible, considering the different treatment strategies. Initial tumour stage and risk features should be considered. In the meta-analysis that included 602 patients from 11 series, advanced cT stage (cT1-2 vs cT3-4) predicted for worse response and local regrowth[30]. Thus, for early-stage tumours treated with CRT or SCRT, the panel recommended the two-step approach adopted by the STAR-TREC trial for RA and determining cCR i.e. 12 weeks and 16-20 weeks after start of treatment, analogous to anal cancer[42]. Following publication of RAPIDO[43] and PRODIGE[44] phase III trials demonstrating improvement in the primary endpoints, disease-related treatment failure (DrTF) and DFS, respectively, the integration of TNT into the management of locally-

advanced rectal cancer is anticipated in updates of treatment guidelines. The panel recommended adaptation of the timepoint of RA for determining cCR according to the TNT duration i.e. 20-38 weeks after treatment start, as currently performed in representative trial examples including OPERA, ACO/ARO/AIO-18.1, GRECCAR12, OPRA and TRIGGER in **Figure 3**. It remains unclear how long it is oncologically safe and meaningful to wait before determining cCR, especially after prolonged TNT. In the RAPIDO, Bahadoer et al. recently raised caution that early response imaging could be advocated to identify patients that might actually progress during preoperative treatment[45]. Close monitoring is important to identify poor responders early to offer immediate surgery. The panel provided these practical recommendations but acknowledged that evidence on optimal timing to determine cCR is far from complete.

The "Amsterdam/Maastricht" criteria were selected for defining cCR and near-cCR[4]. The diagnosis of near-cCR poses a decision challenge. The panel recommended that longer waiting could be considered as performed in several studies[3,5] in case of ncCR, however, this decision should be made also depending on the trial duration. Importantly, based on previous studies[5,22], biopsy was not recommended by the panel, and should not be routinely performed due to risk of being false-negative (e.g. sampling from a fibrotic area) and lack of evidence on its value, especially when DRE, endoscopy and MRI criteria for cCR are fulfilled[1,46]. Indeed, residual cancer cells are often found in the muscularis propia, which can explain the high rate of false negative results of a superficial biopsy[47]. Also, definition of near cCR is difficult as it is not a binary issue that can always be accurately determined by imaging, and depends on the trajectory. Definition of near cCR requires consideration of both regression of lymph nodes with morphological features suspicious for node positivity (round, irregular border and heterogeneous signal) combined with size ≥5 mm [48-51]. LE can be used in the case of ncCR, both for diagnostic and therapeutic purposes[13,52], but can be associated with increased morbidity if completion TME is required[9,12]. The criteria for completion TME after initial LE need to be further elucidated.

Regarding early-stage cancers with an adenomatous component, the largest challenge is the accuracy of diagnosing a residual adenomatous polyp after radiotherapy/CRT of small rectal cancers. Previous data have indicated that these tumours might be suitable for primary treatment with CRT and organ preservation, however, residual adenomatous polyps often include high-grade dysplastic components and should, hence, be removed using full-thickness LE[53,54].

Of note, diagnostic imaging can be notoriously inaccurate at initial diagnosis. Staging is highly relevant in the context of organ preservation as previous studies have indicated increasing cT

stage, tumour volume or, alternatively, tumour length and bowel wall circumferential extend at baseline as the most important predictors of achieving cCR[11,55-57]. Further effort should be made to develop expertise for accurate imaging at diagnosis.

Retrospective and prospective studies have used different methods and follow-up schedules, most of which were designed empirically and extrapolated from oncological guidelines in the operative setting[2-4,6,7,10,58,59]. This was reflected in the large discrepancy of participant votes on follow-up schedule after R1. The panel recommended that follow-up should comprise of CEA, DRE, rectoscopy, pelvic MRI and chest/abdomen CT, and agreed a specific follow-up schedule to avoid inconsistency. Since local regrowth after initial cCR commonly occurs within the first 2-3 years, a period of 3 years of monitoring using all methods was strongly recommended to capture events. Further monitoring was also recommended in the 4th and 5th year as a precaution. Regarding individual methods for organ preservation, a meta-analysis in 602 patients [30] showed that CEA was not a predictor for local regrowth after initial cCR, however, CEA values were missing in 45% of patients, which should be considered when interpreting these findings. Thus, the value of CEA remains unclear and more prospective studies are required to clarify its role. Regarding MRI and endoscopy, analyses have demonstrated their complementary role in determining cCR and predicting local regrowth, although failures of local regrowth detection have been reported[60-63]. The role of CT thorax/abdomen monitoring needs further exploration. We recommend CT imaging every 6-12 months at year 1, and yearly during years 2-5, partly because W&W is not routinely established yet and long-term safety data from randomised studies are missing. In the IWWD, distant metastases were diagnosed in only 8% of 880 patients, mostly during the first 3 years[8]. In a recent systematic review of 17 (mostly retrospective) studies with 1387 patients treated with NOM, the maximum risk for distant metastases was 5.5% in patients with sustained cCR but 23.1% in patients with regrowth after initial cCR, where special caution is needed[64]; similar data were reported by Smith et al[10]. Furthermore, the 5-year incidence of metastases was 28% in bad responders (ypT2–3) after CRT in the GRECCAR2 trial[12] and, thus, special caution is also required in this patient subgroup if LE is explored. Of note, in the updated IWWD report published recently (and after completion of the Delphi process as part of our consensus study), the probability of remaining free from local regrowth for an additional 2 years if a patient had a sustained cCR for 1 year and 3 years was 88.1% and 97.3%, respectively, after a median follow-up of 55.2 months[65]. These data indicated that the intensity of active surveillance if a cCR was sustained could be reduced if they have a sustained cCR within the first 3 years of W&W.

One of the main arguments for exploring NOM is preservation of sphincter and anorectal function. Previous work demonstrated worse anorectal function with major LARS after CRT plus surgery (up to 67%) compared to CRT alone (up to 36%), however, different anorectal function scores have been arbitrarily used[33-35,66]. Despite the lack of evidence from randomised cohorts comparing surgery vs. NOM/LE, the panel recommended that the LARS score[24] is most practical for routine use. The panel acknowledged the limitations of LARS (not validated for organ preservation; only reporting on bowel dysfunction) and recommended that a new PRO specific to organ preservation should be developed.

Although improvement of QoL constituted one of the main arguments for avoiding surgery, randomised evidence on the superiority of (C)RT alone for organ preservation is lacking, other than TREC that demonstrated high levels of organ preservation, with improved QoL after SCRT compared to surgery[14]. Data have mostly been derived from series that used a wide variety of different questionnaires for assessing QoL and PROs, none of which are validated for use in an organ preservation setting[33-35,66]. Thus, the panel agreed for future studies: i) Five QoL and function scales should always be documented; ii) 10 symptomatic toxicity items were selected as highest priority for evaluation; iii) a specific time schedule for measurement; and iv) a new validated questionnaire, or short extension to an existing instrument (e.g. EORTC-QLQ CR29 or LARS) should be developed specifically for organ preservation; designed to capture symptomatic toxicity (bowel, urinary and sexual dysfunction) as well as the impact of more intensive active surveillance on QoL, for use within trials and clinical practice. Importantly, the aspects on QoL and PROs reported here provide the first international consensus and are an important foundation to build upon to harmonise documentation.

Our study has limitations. First, the panel of trialists was selected by design, which could incur bias. Second, the consensus recommendation process was based on online surveys. It was not possible for a face-to-face meeting to discuss discrepancies to take place, but further clarification was possible through email correspondence. Third, although the threshold of 70% required to reach a consensus has been previously used[67-69], it is arbitrary and constitutes a methodological limitation of Delphi surveys[70]. As trial evidence on organ preservation is continuously growing, it is likely that some outcome measures will need adaptation in the future. Thus, the present consensus should serve as guide to further augment rather than fully replace clinical judgment. **Table 5** summarises the key outstanding questions and uncertainties on organ preservation in rectal cancer. Fourth, only health care providers participated in the surveys, whereas other stakeholders (e.g. industry sponsors, patient representatives) were not involved. This was considered essential as organ preservation constitutes a new area of clinical

work and consensus on the highly-complex key outcome measures was needed as a first step. This project will be extended to a wider group with multiple stakeholders including patients in the near future to achieve greater consensus, which will also include development of a new EORTC organ preservation-specific QoL set of items/module. Indeed, patients have partly different perceptions on what they consider relevant in the discussion about their treatment, and differences have been described between the importance assigned by patients and clinicians to clinical and functional outcomes, also in the context of organ preservation[33,71,72].

CONCLUSION

To summarise, to the best of our knowledge, this is the first international expert panel consensus to provide comprehensive and rigorous recommendation on the key outcome measures to be assessed and reported in trials and routine practice of organ preservation in rectal cancer. Implementation of the present consensus has important implications as it will harmonise documenting and reporting organ preservation strategies in rectal cancer to improve interpretation and comparison of new trial findings and standardisation of routine practice.

Conflict of interest: The authors declared no conflicts of interest related to the present manuscript. NPW reports grants from Cancer Research UK outside the submitted work. JJS has served as a clinical advisor for Guardant Health, Inc outside of this work. BM has obtained funding from the NIH, US for research purposes outside the submitted work (Project numbers: 2U19CA021239-35 NCI, NIH/NCI; 1U10 CA180858-01, NIH/NCI). VV declares participation in a company sponsored speaker's bureau and receipt of grants/research supports outside the submitted work: Amgen, Astellas, Astrazeneca, Bayer, Bristol Meyers Squbb, Eisai, Ferrin g, Ipsen, Istitutogentili, Janssen-Cilag, MSD, Merck, Norgine, Novartis farma, Pfizer, Roche, Sanofi, Servier Italia, Varian, Elekta, Vewby;

Author contributions: EF, AA, AG, DSM and CR contributed to literature search and preparation of the Delphi questionnaires. All authors contributed (EF, AA, RGJ, GB, RP, JGA, ER, JJS, CM, FP, MVDV, RBT, ASM, JPG, SPB, MG, RDH, KB, CG, KH, BDM, EL, NW, MAG, VV, MB, AGR, AAA, DSM, CR) to writing and review of the manuscript, and approved the final manuscript. EF and AA are joint first authors. AG, DSM and CR are joint last authors.

Funding acknowledgement: NPW reports grants from Yorkshire Cancer Research, during the conduct of the study.

REFERENCES

1 Beets, G. L., Figueiredo, N. F. & Beets-Tan, R. G. Management of Rectal Cancer Without Radical Resection. *Annu Rev Med* **68**, 169-182, doi:10.1146/annurev-med-062915-021419 (2017).

2 Habr-Gama, A. *et al.* Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg* **240**, 711-717; discussion 717-718 (2004).

3 Habr-Gama, A. *et al.* Organ Preservation in cT2N0 Rectal Cancer After Neoadjuvant Chemoradiation Therapy: The Impact of Radiation Therapy Dose-escalation and Consolidation Chemotherapy. *Ann Surg* **269**, 102-107, doi:10.1097/SLA.0000000000002447 (2019).

4 Maas, M. *et al.* Wait-and-see policy for clinical complete responders after chemoradiation for rectal cancer. *J Clin Oncol* **29**, 4633-4640, doi:10.1200/JCO.2011.37.7176 (2011).

5 Martens, M. H. *et al.* Long-term Outcome of an Organ Preservation Program After Neoadjuvant Treatment for Rectal Cancer. *J Natl Cancer Inst* **108**, doi:10.1093/jnci/djw171 (2016).

6 Renehan, A. G. *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* **17**, 174-183, doi:10.1016/S1470-2045(15)00467-2 (2016).

7 Appelt, A. L. *et al.* High-dose chemoradiotherapy and watchful waiting for distal rectal cancer: a prospective observational study. *Lancet Oncol* **16**, 919-927, doi:10.1016/S1470-2045(15)00120-5 (2015).

8 van der Valk, M. J. M. *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* **391**, 2537-2545, doi:10.1016/S0140-6736(18)31078-X (2018).

9 Rullier, E. *et al.* Organ preservation for rectal cancer (GRECCAR 2): a prospective, randomised, open-label, multicentre, phase 3 trial. *Lancet*, doi:10.1016/S0140-6736(17)31056-5 (2017).

10 Smith, J. J. *et al.* Assessment of a Watch-and-Wait Strategy for Rectal Cancer in Patients With a Complete Response After Neoadjuvant Therapy. *JAMA oncology*, e185896, doi:10.1001/jamaoncol.2018.5896 (2019).

11 Dossa, F., Chesney, T. R., Acuna, S. A. & Baxter, N. N. A watch-and-wait approach for locally advanced rectal cancer after a clinical complete response following neoadjuvant chemoradiation: a systematic review and meta-analysis. *Lancet Gastroenterol Hepatol* **2**, 501-513, doi:10.1016/S2468-1253(17)30074-2 (2017).

12 Rullier, E. *et al.* Organ preservation with chemoradiotherapy plus local excision for rectal cancer: 5-year results of the GRECCAR 2 randomised trial. *Lancet Gastroenterol Hepatol* **5**, 465-474, doi:10.1016/S2468-1253(19)30410-8 (2020).

13 Stijns, R. C. H. *et al.* Long-term Oncological and Functional Outcomes of Chemoradiotherapy Followed by Organ-Sparing Transanal Endoscopic Microsurgery for Distal Rectal Cancer: The CARTS Study. *JAMA Surg* **154**, 47-54, doi:10.1001/jamasurg.2018.3752 (2019).

14 Bach SP, G. A., Brock K, et al. Radical surgery versus organ preservation using short-course radiotherapy followed by transanal endoscopic microsurgery for early-stage rectal cancer (TREC): a randomised, open-label feasibility study. . (Lancet Hepatology Gastroenterology 2020; epub ahead.).

15 Bellera, C. A. *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* **49**, 769-781, doi:10.1016/j.ejca.2012.09.035 (2013).

16 Fokas, E. *et al.* Outcome measures in multimodal rectal cancer trials. *Lancet Oncol* **21**, e252-e264, doi:10.1016/S1470-2045(20)30024-3 (2020).

17 Glynne-Jones, R. *et al.* Rectal cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann Oncol* **28**, iv22-iv40, doi:10.1093/annonc/mdx224 (2017).

18 Benson, A. B. *et al.* NCCN Guidelines Insights: Rectal Cancer, Version 6.2020. *J Natl Compr Canc Netw* **18**, 806-815, doi:10.6004/jnccn.2020.0032 (2020).

19 Wo, J. Y. *et al.* Radiation Therapy for Rectal Cancer: Executive Summary of an ASTRO Clinical Practice Guideline. *Pract Radiat Oncol*, doi:10.1016/j.prro.2020.08.004 (2020).

20 Garcia-Aguilar J, P. S., Kim J, et al. (Preliminary results of the organ preservation of rectal adenocarcinoma (OPRA) trial. J Clin Oncol 38, no. 15_suppl (May 20, 2020) 4008-4008.).

21 Smith, J. J. *et al.* Organ Preservation in Rectal Adenocarcinoma: a phase II randomized controlled trial evaluating 3-year disease-free survival in patients with locally advanced

615 rectal cancer treated with chemoradiation plus induction or consolidation
616 chemotherapy, and total mesorectal excision or nonoperative management. *BMC*
617 *Cancer* **15**, 767, doi:10.1186/s12885-015-1632-z (2015).
618 22 Maas, M. *et al.* Assessment of Clinical Complete Response After Chemoradiation for
619 Rectal Cancer with Digital Rectal Examination, Endoscopy, and MRI: Selection for
620 Organ-Saving Treatment. *Ann Surg Oncol* **22**, 3873-3880, doi:10.1245/s10434-015-
621 4687-9 (2015).
622 23 Jorge, J. M. & Wexner, S. D. Etiology and management of fecal incontinence. *Dis Colon*
623 *Rectum* **36**, 77-97, doi:10.1007/bf02050307 (1993).
624 24 Emmertsen, K. J. & Laurberg, S. Low anterior resection syndrome score: development
625 and validation of a symptom-based scoring system for bowel dysfunction after low
626 anterior resection for rectal cancer. *Ann Surg* **255**, 922-928,
627 doi:10.1097/SLA.0b013e31824f1c21 (2012).
628 25 Temple, L. K. *et al.* The development of a validated instrument to evaluate bowel
629 function after sphincter-preserving surgery for rectal cancer. *Dis Colon Rectum* **48**,
630 1353-1365, doi:10.1007/s10350-004-0942-z (2005).
631 26 Vaizey, C. J., Carapeti, E., Cahill, J. A. & Kamm, M. A. Prospective comparison of
632 faecal incontinence grading systems. *Gut* **44**, 77-80, doi:10.1136/gut.44.1.77 (1999).
633 27 Heald, R. J., Beets, G. & Carvalho, C. Report from a consensus meeting: response to
634 chemoradiotherapy in rectal cancer - predictor of cure and a crucial new choice for the
635 patient: on behalf of the Champalimaud 2014 Faculty for 'Rectal cancer: when NOT to
636 operate'. *Colorectal Dis* **16**, 334-337, doi:10.1111/codi.12627 (2014).
637 28 Buyse, M. *et al.* Statistical evaluation of surrogate endpoints with examples from cancer
638 clinical trials. *Biom J* **58**, 104-132, doi:10.1002/bimj.201400049 (2016).
639 29 Maas, M. *et al.* Long-term outcome in patients with a pathological complete response
640 after chemoradiation for rectal cancer: a pooled analysis of individual patient data.
641 *Lancet Oncol* **11**, 835-844, doi:10.1016/S1470-2045(10)70172-8 (2010).
642 30 Chadi, S. A. *et al.* Factors affecting local regrowth after watch and wait for patients with
643 a clinical complete response following chemoradiotherapy in rectal cancer (InterCoRe
644 consortium): an individual participant data meta-analysis. *Lancet Gastroenterol*
645 *Hepatol* **3**, 825-836, doi:10.1016/S2468-1253(18)30301-7 (2018).
646 31 Battersby, N. J. *et al.* A rectal cancer feasibility study with an embedded phase III trial
647 design assessing magnetic resonance tumour regression grade (mrTRG) as a novel
648 biomarker to stratify management by good and poor response to chemoradiotherapy

(TRIGGER): study protocol for a randomised controlled trial. *Trials* **18**, 394, doi:10.1186/s13063-017-2085-2 (2017).

32 Saad, E. D., Paoletti, X., Burzykowski, T. & Buyse, M. Precision medicine needs randomized clinical trials. *Nat Rev Clin Oncol* **14**, 317-323, doi:10.1038/nrclinonc.2017.8 (2017).

33 van der Valk, M. J. M. *et al.* Importance of patient reported and clinical outcomes for patients with locally advanced rectal cancer and their treating physicians. Do clinicians know what patients want? *Eur J Surg Oncol* **46**, 1634-1641, doi:10.1016/j.ejso.2020.04.014 (2020).

34 van der Sande, M. E. *et al.* Impact of radiotherapy on anorectal function in patients with rectal cancer following a watch and wait programme. *Radiother Oncol* **132**, 79-84, doi:10.1016/j.radonc.2018.11.017 (2019).

35 Dizdarevic, E. *et al.* Long-Term Patient-Reported Outcomes After High-Dose Chemoradiation Therapy for Nonsurgical Management of Distal Rectal Cancer. *Int J Radiat Oncol Biol Phys* **106**, 556-563, doi:10.1016/j.ijrobp.2019.10.046 (2020).

36 Peltrini, R., Sacco, M., Luglio, G. & Bucci, L. Local excision following chemoradiotherapy in T2-T3 rectal cancer: current status and critical appraisal. *Updates Surg* **72**, 29-37, doi:10.1007/s13304-019-00689-2 (2020).

37 Smith, F. M. *et al.* Local Excision Techniques for Rectal Cancer After Neoadjuvant Chemoradiotherapy: What Are We Doing? *Dis Colon Rectum* **60**, 228-239, doi:10.1097/DCR.0000000000000749 (2017).

38 Arezzo, A. *et al.* Individual participant data pooled-analysis of risk factors for recurrence after neoadjuvant radiotherapy and transanal local excision of rectal cancer: the PARTTLE study. *Tech Coloproctol* **23**, 831-842, doi:10.1007/s10151-019-02049-z (2019).

39 Sloothaak, D. A. *et al.* Optimal time interval between neoadjuvant chemoradiotherapy and surgery for rectal cancer. *Br J Surg* **100**, 933-939, doi:10.1002/bjs.9112 (2013).

40 Garcia-Aguilar, J. *et al.* Effect of adding mFOLFOX6 after neoadjuvant chemoradiation in locally advanced rectal cancer: a multicentre, phase 2 trial. *Lancet Oncol* **16**, 957-966, doi:10.1016/S1470-2045(15)00004-2 (2015).

41 Fokas, E. *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*, JCO1900308, doi:10.1200/JCO.19.00308 (2019).

42 Glynne-Jones, R. *et al.* Best time to assess complete clinical response after chemoradiotherapy in squamous cell carcinoma of the anus (ACT II): a post-hoc analysis of randomised controlled phase 3 trial. *Lancet Oncol* **18**, 347-356, doi:10.1016/S1470-2045(17)30071-2 (2017).

43 Bahadoer, R. R. *et al.* Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *Lancet Oncol*, doi:10.1016/S1470-2045(20)30555-6 (2020).

44 Conroy, T. *et al.* Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): a multicentre, randomised, open-label, phase 3 trial. *Lancet Oncol*, doi:10.1016/S1470-2045(21)00079-6 (2021).

45 Bahadoer, R. R. *et al.* Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): a randomised, open-label, phase 3 trial. *Lancet Oncol* **22**, 29-42, doi:10.1016/S1470-2045(20)30555-6 (2021).

46 Sun, W. *et al.* Controversies in Rectal Cancer Treatment and Management. *Am Soc Clin Oncol Educ Book* **40**, 1-11, doi:10.1200/EDBK_279871 (2020).

47 Duldulao, M. P. *et al.* Distribution of residual cancer cells in the bowel wall after neoadjuvant chemoradiation in patients with rectal cancer. *Dis Colon Rectum* **56**, 142-149, doi:10.1097/DCR.0b013e31827541e2 (2013).

48 Kim, J. H., Beets, G. L., Kim, M. J., Kessels, A. G. & Beets-Tan, R. G. High-resolution MR imaging for nodal staging in rectal cancer: are there any criteria in addition to the size? *Eur J Radiol* **52**, 78-83, doi:10.1016/j.ejrad.2003.12.005 (2004).

49 Heijnen, L. A. *et al.* Nodal staging in rectal cancer: why is restaging after chemoradiation more accurate than primary nodal staging? *Int J Colorectal Dis* **31**, 1157-1162, doi:10.1007/s00384-016-2576-8 (2016).

50 van Heeswijk, M. M. *et al.* DWI for Assessment of Rectal Cancer Nodes After Chemoradiotherapy: Is the Absence of Nodes at DWI Proof of a Negative Nodal Status? *AJR Am J Roentgenol* **208**, W79-W84, doi:10.2214/AJR.16.17117 (2017).

51 Perez, R. O. *et al.* Lymph node size in rectal cancer following neoadjuvant chemoradiation--can we rely on radiologic nodal staging after chemoradiation? *Dis Colon Rectum* **52**, 1278-1284, doi:10.1007/DCR.0b013e3181a0af4b (2009).

52 Garcia-Aguilar, J. *et al.* Organ preservation for clinical T2N0 distal rectal cancer using neoadjuvant chemoradiotherapy and local excision (ACOSOG Z6041): results of an open-label, single-arm, multi-institutional, phase 2 trial. *Lancet Oncol* **16**, 1537-1546, doi:10.1016/S1470-2045(15)00215-6 (2015).

53 Rupinski, M. *et al.* Watch and wait policy after preoperative radiotherapy for rectal cancer; management of residual lesions that appear clinically benign. *Eur J Surg Oncol* **42**, 288-296, doi:10.1016/j.ejso.2015.09.022 (2016).

54 Habr-Gama, A. *et al.* Management of adenomas within the area of rectal cancer that develop complete pathological response. *Int J Colorectal Dis* **30**, 1285-1287, doi:10.1007/s00384-015-2326-3 (2015).

55 Rijkmans, E. C. *et al.* Predictive factors for response and toxicity after brachytherapy for rectal cancer; results from the HERBERT study. *Radiother Oncol* **133**, 176-182, doi:10.1016/j.radonc.2019.01.034 (2019).

56 van Stiphout, R. G. *et al.* Nomogram predicting response after chemoradiotherapy in rectal cancer using sequential PETCT imaging: a multicentric prospective study with external validation. *Radiother Oncol* **113**, 215-222, doi:10.1016/j.radonc.2014.11.002 (2014).

57 Das, P. *et al.* Predictors of tumor response and downstaging in patients who receive preoperative chemoradiation for rectal cancer. *Cancer* **109**, 1750-1755, doi:10.1002/cncr.22625 (2007).

58 Garcia-Aguilar, J. *et al.* Organ preservation for clinical T2N0 distal rectal cancer using neoadjuvant chemoradiotherapy and local excision (ACOSOG Z6041): results of an open-label, single-arm, multi-institutional, phase 2 trial. *Lancet Oncol* **16**, 1537-1546, doi:10.1016/S1470-2045(15)00215-6 (2015).

59 Gerard, J. P. *et al.* Planned organ preservation for early T2-3 rectal adenocarcinoma: A French, multicentre study. *Eur J Cancer* **108**, 1-16, doi:10.1016/j.ejca.2018.11.022 (2019).

60 Patel, U. B., Blomqvist, L., Chau, I., Nicholls, J. & Brown, G. Session 3: Beyond TME and radiotherapy MRI evaluation of rectal cancer treatment response. *Colorectal Dis* **20 Suppl 1**, 76-81, doi:10.1111/codi.14084 (2018).

61 Lambregts, D. M. J. *et al.* Long-term imaging characteristics of clinical complete responders during watch-and-wait for rectal cancer-an evaluation of over 1500 MRIs. *Eur Radiol* **30**, 272-280, doi:10.1007/s00330-019-06396-1 (2020).

62 Lambregts, D. M. J. *et al.* A Pattern-Based Approach Combining Tumor Morphology on MRI With Distinct Signal Patterns on Diffusion-Weighted Imaging to Assess Response of Rectal Tumors After Chemoradiotherapy. *Dis Colon Rectum* **61**, 328-337, doi:10.1097/DCR.0000000000000915 (2018).

63 van Griethuysen, J. J. M. *et al.* Radiomics performs comparable to morphologic assessment by expert radiologists for prediction of response to neoadjuvant chemoradiotherapy on baseline staging MRI in rectal cancer. *Abdom Radiol (NY)* **45**, 632-643, doi:10.1007/s00261-019-02321-8 (2020).

64 Socha, J., Kepka, L., Michalski, W., Paciorek, K. & Bujko, K. The risk of distant metastases in rectal cancer managed by a watch-and-wait strategy - A systematic review and meta-analysis. *Radiother Oncol* **144**, 1-6, doi:10.1016/j.radonc.2019.10.009 (2020).

65 Fernandez, L. M. *et al.* Conditional recurrence-free survival of clinical complete responders managed by watch and wait after neoadjuvant chemoradiotherapy for rectal cancer in the International Watch & Wait Database: a retrospective, international, multicentre registry study. *Lancet Oncol* **22**, 43-50, doi:10.1016/S1470-2045(20)30557-X (2021).

66 Hupkens, B. J. P. *et al.* Quality of Life in Rectal Cancer Patients After Chemoradiation: Watch-and-Wait Policy Versus Standard Resection - A Matched-Controlled Study. *Dis Colon Rectum* **60**, 1032-1040, doi:10.1097/DCR.0000000000000862 (2017).

67 Fish, 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* **3**, 865-873, doi:10.1016/S2468-1253(18)30264-4 (2018).

68 Hui, D. *et al.* Referral criteria for outpatient specialty palliative cancer care: an international consensus. *Lancet Oncol* **17**, e552-e559, doi:10.1016/S1470-2045(16)30577-0 (2016).

69 Hasson, F., Keeney, S. & McKenna, H. Research guidelines for the Delphi survey technique. *J Adv Nurs* **32**, 1008-1015 (2000).

70 Kirkham, J. J. *et al.* Core Outcome Set-STAndards for Development: The COS-STAD recommendations. *PLoS Med* **14**, e1002447, doi:10.1371/journal.pmed.1002447 (2017).

71 Kunneman, M., Pieterse, A. H., Stiggelbout, A. M. & Marijnen, C. A. Which benefits and harms of preoperative radiotherapy should be addressed? A Delphi consensus study

among rectal cancer patients and radiation oncologists. *Radiother Oncol* **114**, 212-217, doi:10.1016/j.radonc.2014.11.034 (2015).

72 Gani, C. *et al.* Organ Preservation in Rectal Cancer: The Patients' Perspective. *Front Oncol* **9**, 318, doi:10.3389/fonc.2019.00318 (2019).

FIGURE LEGENDS

Figure 1. Illustration of the flowchart of article selection process. Seven key outcome measures of organ preservation strategies in rectal cancer were identified following a thorough literature search.

Figure 2. Summarized overview of the Delphi process

Figure 3. Illustration of the timepoints of response assessment (RA) to determine clinical complete response (cCR), and corresponding primary endpoints according to the different trial phase and design in representative examples of organ preservation trials. The different preoperative/definitive treatment options that are characterised by variable length and time to RA and decision on organ preservation *vs* total mesorectal excision surgery appear below the x-axis. Examples of corresponding clinical trials with the TNM stage and treatment arms are shown on the left side marked with dark blue colour (also summarised in **Table 2** that, similarly to the figure, only includes randomised studies). The timepoint of RA and, hence, determining cCR in the different trials is marked with orange colour. The primary endpoint of the trials is shown on the right side with light blue colour. The advent of total neoadjuvant treatment with highly variable duration has added to the complexity of deciding about the optimal timepoint of RA. Abbreviations: Txt, treatment; NOM, non-operative management; LE, local excision; TME, total mesorectal excision; cTNM, clinical tumour/node/metastasis staging; MRI, magnetic resonance imaging; DRE, digital rectal examination; CEA, carcinoembryonic antigen; AV, anal verge; SCRT, short-course radiotherapy; CRT, chemoradiotherapy; SIB, simultaneous integrated boost of radiotherapy; RA, response assessment; OP, organ preservation; DFS, disease-free survival.

849

850 **Table 1.** Definitions of clinical endpoints for organ preservation strategies in rectal cancer

Clinical endpoint	Definition
Organ preservation	Rectum intact (i.e. no radical TME-surgery), no locoregional regrowth unless amenable to limited, curative (R0) salvage surgery by LE, no permanent stoma (i.e., never reversed protective stoma, or stoma due to toxicity or poor functional outcome).
Clinical complete response (cCR)[$]	DRE and Rectoscopy: no palpable tumour, no residual tumour or only a small residual erythematous ulcer or scar; MRI[§]: Substantial downsizing with no residual tumour or residual fibrosis only (with low signal on diffusion-weighted imaging), sometimes associated with residual wall thickening due to edema; no suspicious lymph nodes Endoscopic biopsy: not mandatory to define cCR; biopsy should not be performed, especially if the DRE, rectoscopy and MRI criteria for cCR are fulfilled
Near cCR (ncCR)	DRE and Rectoscopy: Small and smooth regular irregularity; Residual ulcer, or small mucosal nodules or minor mucosal abnormalities, with mild persisting erythema of the scar MRI: Regression of lymph nodes with no malignant enhancement features but size >5 mm
Poor response	Palpable tumour mass and visible macroscopic tumour and/or lack of regression of involved lymph nodes (i.e. patients that do not fulfill the criteria for either cCR or ncCR)
Locoregional regrowth	An event involving either the bowel wall, mesorectum and/or pelvic organs that occurs after initial cCR and W&W
Local regrowth	An event involving the bowel wall only that occurs after initial cCR and W&W
Locoregional recurrence	An event involving either the bowel wall, mesorectum and/or pelvic organs that occurs after LE or TME
Local recurrence	An event involving the bowel wall only that occurs after LE or TME
TME-free DFS[&]	Time from randomisation to one of the following events: radical TME surgery for non-complete response at re-staging, any locoregional regrowth after initial cCR requiring salvage-TME, any locoregional recurrence after LE or no-salvageable regrowth (a regrowth that cannot be removed with an R0 resection), distant metastasis or death (all cause), whichever occurs first
Organ preservation-adapted DFS*	Time from randomisation to one of the following events: No resection of primary tumour due to local progression or patient unfit for surgery, non-radical resection of primary tumour (R2-resection), locoregional recurrence after R0/1 resection of the primary tumour, non-salvageable local regrowth in case of NOM management (no operation or R2 salvage resection), any distant metastatic disease before, at, or after surgery or NOM management, second primary colorectal cancer, Second primary, other cancer, treatment-related death, death from same cancer, death from other cancer, non-cancer related death

Abbreviations: TME, total mesorectal excision; DRE, digital rectal examination; cCR, clinical complete response; W&W, watch and wait; LE, local excision; DFS, disease-free survival

[$]All criteria of DRE, rectoscopy and MRI should be fulfilled to define cCR

[§]Gadolinium contrast medium is no longer compulsory for MRI conducted to define clinical complete response

[&]Consensus was not reached for the definition of TME-free DFS that was provided separately by the primary investigator of the OPRA trial, JGA.

*If a salvage operation for the local regrowth is performed in curative intent (R0/1), it should not count as an event. If, however, no operation, or only a R2 resection is possible, or there is a recurrence after salvage surgery, this should count as an event.

Table 2. Summary of randomised clinical trials of organ preservation in rectal cancer showing the large variability in the timepoints of response assessment (RA) used to determine clinical complete response (cCR) as well as the primary endpoints used among the different trials

Trial and TNM staging*, &	N	Treatment schedule	Timepoint of response assessment (RA)	Primary endpoint
TREC, phase 2[14] cT1-23N0, maximum diameter ≤30 mm (ISRCTN 14422743)	55	TME surgery vs SCRT followed by TEM	9-11 weeks after treatment start	Recruitment rate at 12, 18 and 24 months
STAR TREC, phase 3 part (NCT02945566) cT1-T3bN0, ≤10 cm AV	460	TME/LE surgery ***vs*** CRT followed by NOM/LE vs SCRT followed by NOM/LE (if cCR: NOM; if PR: TEM; if poor response: TME)	12 and 20 weeks after treatment start	30-month organ preservation rate
WW3, phase 2 (NCT04095299) cT1-T3bN0, ≤10 cm AV	111	CRT ***vs*** CRT with SIB (if cCR: NOM or LE; if partial response: TME)	16 weeks after treatment start	2-year organ preservation
OPERA, phase 3 (NCT02505750) cT2-T3bN0-1, ≤10 cm AV	236	CRT followed by EBRT boost vs CRT followed by brachytherapy boost (if cCR: NOM or LE; if PR: TME)	14 and 20-24 weeks after treatment start	3-year organ preservation
HERBERT-II, phase 3 (NL7795), elderly and frail with cT1-3N0-1, ≤10 cm AV	106	EBRT vs EBRT plus brachytherapy boost	26 weeks after treatment end	cCR rate at 26 weeks
GRECCAR12, phase 3 (NCT02514278) cT2-T3N0-1, ≤10 cm AV	218	mFOLFIRINOX followed by CRT vs CRT (if good response: LE; if poor response: TME)	24 weeks after treatment start	12-month organ preservation
ACO/ARO/AIO-18.1, phase 3, (NCT04246684) cT3c-T4N0-2, ≤12 cm AV	702	SCRT followed by consolidation FOLFOX chemotherapy and TME surgery (or NOM if cCR) ***vs*** CRT followed by consolidation FOLFOX chemotherapy and TME surgery (or NOM if cCR)	24 weeks after treatment start	3-year organ preservation
OPRA[20,21], phase 2 (NCT02008656) cT3-T4N0-2, ≤6 cm AV	300	Induction mFOLFOX6 chemotherapy followed by CRT and surgery/NOM ***vs*** CRT followed by consolidation mFOLFOX6 chemotherapy and surgery/NOM	34-38 weeks after treatment start	3-year DFS
TRIGGER[31], phase 2/3, (NCT02704520) cT3c-T4N0-2, ≤15 cm AV	90	CRT followed by surgery and adjuvant CAPOX/FOLFOX ***vs*** CRT followed by either NOM (mrTRG I-II) or CAPOX/FOLFOX (mrTRG III-IV) and restaging with subsequent NOM or surgery (depending on mrTRG at restaging)	12, 24 and 36-38 weeks after treatment start	Recruitment rate (phase 2); 3-year DFS (phase 3)
Brazilian[§], phase 3 (NCT02052921) cT3-T4N0-2, ≤10 cm AV	150	CRT followed by W&W vs 5-FU CRT followed by TME surgery after achieving cCR at 12 weeks post CRT	12 weeks after treatment start	3-year DFS
TESAR, phase 2 (NCT02371304) pT1-2cN0, ≤10 cm AV	302	TME surgery vs LE followed by CRT	n.a.	3-year LRR
MORPHEUS, phase 2 (NCT03051464) cT2-T3bN0, ≤10 cm AV	40	CRT followed by EBRT boost vs CRT followed by brachytherapy boost (if cCR: NOM; if PR: TME)	14 weeks after treatment start	2-year organ preservation
TESS, phase 2, (NCT03840239) cT3-4aN0-2, ≤5cm AV	168	Induction CAPOX followed by CRT vs CRT (if cCR: NOM; if PR: LE or TEM; if poor response: TME)	20-24 weeks after treatment start	Sphincter preservation (stoma absence) at 18 months
APHRODITE, phase 2 (ISRCTN16158514) cT1-T3bN0, ≤10 cm AV	104	CRT vs CRT with SIB (if cCR: NOM)	24 weeks after treatment start	cCR rate at 6 months

GRECCAR2[9,12], phase 3 (NCT00427375) cT2-3N0–1, ≤5 cm AV maximum initial size 4 cm residual tumour ≤2 cm	186	CRT followed by local excision vs preoperative CRT followed by TME surgery	12-14 weeks after treatment start	2-year Composite endpoint
ELRRvsLTME, phase 3 (NCT01609504) cT2N0, ≤6 cm AV	100	CRT followed by local excision vs CRT followed by TME surgery	n.a.	Local and distant recurrence (timepoint unspecified)

Abbreviations: CRT, chemoradiotherapy; TME, total mesorectal excision; DFS, disease-free survival; NOM, non-operative management; cCR, clinical complete response; LE, local excision; SCRT, short-course radiotherapy; W&W, watch and wait; TdrTF, Time to Disease-related Treatment Failure; LRR, locoregional recurrence, mrTRG, magnetic resonance-based tumour regression grading; TNT, total neoadjuvant treatment; CAPOX, capecitabine/oxaliplatin;

*Only randomised studies on organ preservation were included in this table.

[&]Tumour location, especially for rectal cancers close to the anal sphincter where often the only surgical option is abdominoperineal resection with permanent stoma, can influence the use of CRT for early-stage disease to achieve organ preservation, as reflected in many trials that included patients with cT2 rectal cancer.

[§]The Brazilian trial was closed prematurely (May 2020) due to poor patient accrual. This was the first clinical trial to randomise patients with cCR after preoperative chemoradiotherapy to W&W *vs* surgery, and used DFS as a primary endpoint.

Table 3. Panel consensus recommendation on the timepoint of response assessment (RA) to determine clinical complete response (cCR) according to the treatment design and duration

Treatment design	**Time point of response assessment (RA)**
Standard SCRT (duration: 5 days) or CRT (duration: ~6 weeks) for early-stage tumours,	A two-step approach is recommended, i.e. measurement at 12 weeks from the start of treatment and then, in case of ncCR at initial assessment, again at 16-20 weeks, should be used to determine cCR (e.g. STAR-TREC trial: NCT02945566).
CRT followed by brachytherapy (duration: 12 weeks)	cCR should be determined at 14 weeks and, in case of ncCR at initial assessment, at 20-24 weeks, after start of treatment (e.g. OPERA trial: NCT02505750).
TNT with CRT and either induction or consolidation chemotherapy (duration: 16-20 weeks)	cCR should be determined at 24 weeks after start of treatment (e.g. GRECCAR12 trial: NCT02514278 and ACO/ARO/AIO-18.1 trial: NCT04246684, trials, respectively).
TNT with SCRT/CRT followed by **prolonged** consolidation chemotherapy (duration: 26 and 34 weeks, as in OPRA and TRIGGER trials, respectively),	cCR should be determined at 34-38 weeks after start of treatment (e.g. OPRA trial[20] and TRIGGER trial: NCT02704520[31]).

Abbreviations: SCRT, short-course radiotherapy; CRT, chemoradiotherapy; TNT, total neoadjuvant treatment; cCR, clinical complete response; ncCR, near cCR;

Table 4. Consented follow-up methods and schedule for organ preservation strategy

Year	CEA	DRE	Endoscopy	MRI pelvis	Computed tomograpy chest/abdomen
1	3×	3-4×	3-4×	3-4×	1-2×
2	3×	3-4×	3-4×	3-4×	1×
3	3×	2×	2×	2×	1×
4	2×	2×	2×	2×	1×
5	2×	2×	2×	2×	1×

Abbreviations: CEA, carcinoembryonic antigen; DRE, digital rectal examination; MRI, magnetic resonance imaging;

First follow-up assessment commonly occurs 6-8 weeks upon completion of preoperative / definitive treatment

Table 5. Summary of key outstanding questions and current uncertainties on organ preservation in rectal cancer

Key outstanding questions
1. Which criteria should we use to include patient in studies of organ preservation?
2. Can modern technology methods (e.g. artificial intelligence and neural networks) help to improve accuracy of imaging at initial diagnosis of rectal cancer, and to assess tumor response to treatment?
3. How long is it oncologically safe and meaningful to wait to assess tumor response before determining cCR, especially after prolonged TNT?
4. What is the role of LE as primary treatment, and for selected patients with good response after CRT?
5. What is the optimal time for LE in the context of tumor response (cCR vs near cCR vs residual disease)?
6. Which criteria should we use to advocate LE for organ preservation?
7. What is the optimal surgical method to manage regrowth after initial cCR?
8. Can we define robust selection criteria to safely reduce the intensity of follow-up imaging in patients with cCR?
9. What is the long-term impact of the different strategies explored for “intended” organ preservation (selective CRT with LE; RT dose escalation; TNT etc) on QoL, function as well as short and long term toxicity?
10. Which items and function scales should be included in a PRO designed specifically for organ preservation?
11. Can liquid biopsy biomarkers (e.g. CEA, circulating or free DNA) be used to predict cCR and tumour regrowth after initial cCR to tailor treatment?