Defining near-complete response following (chemo)radiotherapy for rectal cancer: systematic review
Custers, P.A.; Geubels, B.M.; Beets, G.L.; Lambregts, D.M.J.; Leerdam, M.E. van; Triest, B. van; Maas, M.

Citation

Version: Publisher's Version
License: Leiden University Non-exclusive license
Downloaded from: https://hdl.handle.net/1887/3492270

Note: To cite this publication please use the final published version (if applicable).
Defining near-complete response following (chemo) radiotherapy for rectal cancer: systematic review

Petra A. Custers1,2, Barbara M. Geubels1,2,3, Geerard L. Beets1,2, Doenja M. J. Lambregts4, Monique E. van Leerda5,6, Baukelien van Triest7 and Monique Maas4,*

1Department of Surgery, Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam, the Netherlands
2GROW School for Oncology and Reproduction, Maastricht University, Maastricht, the Netherlands
3Department of Surgery, Catharina Hospital, Eindhoven, the Netherlands
4Department of Radiology, Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam, the Netherlands
5Department of Gastroenterology and Hepatology, Leiden University Medical Centre, Leiden, the Netherlands
6Department of Radiation Oncology, Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam, the Netherlands
7Department of Radiation Oncology, Netherlands Cancer Institute, Amsterdam, the Netherlands
8Department of Radiation Oncology, Netherlands Cancer Institute—Antoni van Leeuwenhoek, Amsterdam, the Netherlands
9Department of Surgery, Netherlands Cancer Institute

Abstract

Background: A uniform definition of a clinical near-complete response (near-CR) after neoadjuvant (chemo)radiotherapy for rectal cancer is lacking. A clear definition is necessary for uniformity in clinical practice and trial enrolment for organ-preserving treatments. This review aimed to provide an overview of the terminology, criteria, and features used in the literature to define a near-CR.

Methods: A systematic review was performed based on the PRISMA statement. PubMed and Embase were searched up to May 2021 to identify the terminology, criteria, and features used to define a near-CR after (chemo)radiotherapy for rectal cancer. Studies with no clear cut-off point between a cCR and near-CR, studies using Response Evaluation Criteria In Solid Tumours, and studies including only complete responders were excluded.

Results: A total of 1876 articles were found, of which 23 were included. Patients were managed by watchful waiting and/or additional local treatment in 11 and 17 of 23 studies respectively. Response evaluation included digital rectal examination (DRE) and/or endoscopy with MRI in 18 studies. The majority of studies used the term ‘near-complete response’. In most studies, minor irregularities or a smooth induration with DRE and a small flat ulcer on endoscopy were considered to indicate a near-CR. On MRI, five studies used features (obvious downstaging with or without heterogeneous/irregular fibrosis on T2-weighted MRI or small spot of high signal on diffusion-weighted imaging), five studies used TNM criteria (ycT2), and four used magnetic resonance tumour regression grade (mTRG) (mTRG2) to describe a near-CR.

Conclusion: The terminology, criteria, and features used to describe a near-CR vary substantially, which can partly be explained by the different treatment strategies patients are selected for (watchful waiting or additional local treatment). A reproducible definition of near-CR is required.

Introduction

In the management of rectal cancer, organ preserving treatment strategies, such as watchful waiting and additional local treatment strategies, have emerged as viable and successful treatment options1–4. Organ preservation has several potential benefits compared with total mesorectal excision. Surgical morbidity and mortality can be avoided (including a stoma), and organ preservation leads to better quality of life and functional outcome than surgery4. For patients with a clinical complete response (cCR) after neoadjuvant chemoradiotherapy, organ preservation does not compromise oncological results, as local regrowth (occurring in 25 per cent) can usually be curatively treated without a negative impact on long-term outcome4,5.

To select patients for watchful waiting or organ preservation, a combination of digital rectal examination (DRE), endoscopy, and T2-weighted (T2W) MRI combined with diffusion-weighted MRI (DWI) appears to be most accurate6. Using these modalities, the criteria used to select patients with a cCR for watchful waiting are well described6,7. These criteria, as outlined in the initial publications, were quite strict, aiming to avoid undertreatment. A drawback of these strict criteria was an underestimation of response in a substantial part of patients, thereby denying them the chance of organ preservation. The concept of a near-complete response (near-CR) was introduced to address this issue8-10. In patients with a near-CR, significant downsizing of the tumour has typically occurred, although not all criteria of a cCR are met. Previous studies8,9 have shown that these patients can progress to a cCR after an additional waiting interval of 6–12 weeks. To increase the possibility of organ preservation, patients with a near-CR may also receive additional local treatment, such as a local excision or a radiotherapy boost11-13.

To ensure safety in clinical practice and consistency in studies, a clear definition is necessary to identify patients with a near-CR. However, in contrast to the widespread consensus on the criteria for a cCR, a variety of features and criteria have been reported to define a near-CR. Also, the terminology used for the good response that enables organ-preservation is variable. A systematic review
was undertaken to identify the various features and criteria used in the literature to describe a near-CR.

**Methods**

Two independent reviewers searched PubMed and Embase databases from inception to May 2021 to identify the criteria and features used to define a near-CR. As the term ‘near-complete response’ is not used universally, the search aimed to identify the criteria and features of a near-CR or synonyms and similar terms, such as ‘good response’, which described a response indicating that patients are potentially eligible for organ-preserving treatments. The Medical Subject Heading (MESH) terms and free search terms used in the search are provided in Tables S1 and S2.

Studies were included if they met the following criteria: included patients diagnosed with non-metastasized rectal cancer; patients were treated with neoadjuvant therapy; clinical response evaluation (restaging) was performed following neoadjuvant therapy; either the term ‘near-complete response’ was used or patients were selected for organ-preserving treatments based on a good response, although not a complete response; and patients were treated with curative intent. Studies were excluded where: only patients with a cCR were included; no clear cut-off point was provided between a cCR and near-CR; and the clinical response was based on Response Evaluation Criteria In Solid Tumours criteria. Additionally, conference abstracts, case reports, reviews, and studies not reporting in English were excluded. To provide a complete overview of the terminology and definitions used for a near-CR in the literature, both prospective and retrospective studies were included in this review. Studies with potentially overlapping cohorts were also included as the terminology and definitions of a near-CR may vary between these studies. When articles reported on exactly the same patient cohort, the most recently published article was included.

Potentially eligible studies were selected based on title and abstract. Full-text copies of these studies were reviewed to select the studies that met the inclusion criteria. In addition, references of all selected studies were checked for eligibility. Disagreement was resolved by a third reviewer. Data from all selected studies were extracted independently by the two reviewers. The following data were extracted: study design; number of patients included; tumour and patient characteristics; treatment details; details of the timing of, and modalities used, for response evaluation; duration of follow-up; response classification; and criteria and features associated with a near-CR. No automated tools were used in the selection or data extraction process.

The methodological quality of the included studies was assessed to gain insight into the overall methodological quality of studies reporting on a near-complete response or organ-preserving treatments. To assess methodological quality, a checklist was composed, based mainly on the STROBE checklist, complemented by some items from the CONSORT and STARD checklists to fit the present research question. Items concerning details of randomization, blinding, index test, and reference standard were
Table 1 Summary of findings: criteria and features defining a near-complete response by modality used

<table>
<thead>
<tr>
<th>Features defining a near-complete response</th>
</tr>
</thead>
<tbody>
<tr>
<td>Digital rectal examination</td>
</tr>
<tr>
<td>- Normal digital rectal examination</td>
</tr>
<tr>
<td>- Firm rectal wall</td>
</tr>
<tr>
<td>- No palpable mass/no evidence of mucosal malignancy</td>
</tr>
<tr>
<td>- Smooth induration</td>
</tr>
<tr>
<td>- Minor mucosal abnormalities/necrosis</td>
</tr>
<tr>
<td>- Soft superficial irregularity (&lt; 2 cm)</td>
</tr>
<tr>
<td>- Decrease in tumour size &gt; 50%</td>
</tr>
<tr>
<td>Endoscopy</td>
</tr>
<tr>
<td>- No mucosal abnormality</td>
</tr>
<tr>
<td>- No evidence of mucosal malignancy</td>
</tr>
<tr>
<td>- Residual scar ≤ 3 cm in diameter/whitening of mucosa</td>
</tr>
<tr>
<td>- Mild persistent erythema of scar</td>
</tr>
<tr>
<td>- Small supple/firm nodularity or irregularity</td>
</tr>
<tr>
<td>- Superficial ulcer (&lt; 2 cm or &lt; 3 cm)</td>
</tr>
<tr>
<td>- Residual lesions (&lt; 3 cm)</td>
</tr>
<tr>
<td>- Residual tumour (&lt; 2 cm or &lt; 3 cm)</td>
</tr>
<tr>
<td>- Dysplasia at histopathology</td>
</tr>
<tr>
<td>MRI</td>
</tr>
<tr>
<td>- Total mpMRI score 4–5, based on sum of post-CRT T2W-MRI, DWI, and postcontrast scores</td>
</tr>
<tr>
<td>- Obvious downstaging with or without residual fibrosis, but with heterogeneous or irregular aspect</td>
</tr>
<tr>
<td>- Low or intermediate residual signal</td>
</tr>
<tr>
<td>- Small residual lesion with uncertain tumour viability</td>
</tr>
<tr>
<td>- mrTRG1–2 or mrTRG2</td>
</tr>
<tr>
<td>- ycT1–2 or ycT2</td>
</tr>
<tr>
<td>- Small focal area of high signal</td>
</tr>
<tr>
<td>- Significant regression of signal</td>
</tr>
<tr>
<td>- Minimal or low residual signal</td>
</tr>
<tr>
<td>Lymph nodes</td>
</tr>
<tr>
<td>- Absence of a positive regional lymph node</td>
</tr>
<tr>
<td>- Partial regression of lymph nodes</td>
</tr>
<tr>
<td>- Obvious downstaging of lymph nodes but remaining node(s) ≥ 5 mm</td>
</tr>
<tr>
<td>PET–CT</td>
</tr>
<tr>
<td>- Restricted to bowel wall (no evidence of perirectal fat invasion—ycT1–2)</td>
</tr>
<tr>
<td>- Small residual lesion with uncertain tumour viability</td>
</tr>
<tr>
<td>- No evidence of nodal metastases (ycN0)</td>
</tr>
<tr>
<td>- No evidence of nodal metastases (ycN0)</td>
</tr>
</tbody>
</table>

mpMRI, multiparametric MRI; CRT, chemoradiotherapy; T2W-MRI, T2-weighted MRI; DWI, diffusion-weighted MRI; mrTRG, magnetic resonance tumour regression grade; ADC, apparent diffusion coefficient.

found not to be applicable. Results are reported according to the PRISMA statement (Table S3)\textsuperscript{17,18}.

Results

Literature search

A total of 3443 records were identified in PubMed and Embase. After removal of duplicates and exclusions, 23 articles\textsuperscript{5,8–11,19–36} with a total of 1845 patients were included. Figure 1 details the process and reasons for exclusion.

The quality assessment of the 23 included articles is summarized in Table S4. Seventeen described therapeutic studies, of which 10 had a prospective study design. All 23 studies were of adequate quality with regard to the inclusion and exclusion criteria for patients, and duration of follow-up. However, most studies were of insufficient quality with regard to the required sample size and the use of a comparator group. Overall, the included studies reporting on a near-CR or organ-preserving treatments were of moderate to good methodological quality. No studies were excluded based on methodological quality. Study characteristics are provided in Table S5.

Treatment

All 23 studies included patients treated with a long course of (chemo)radiotherapy (45–56 Gy). Four studies\textsuperscript{8,9,25,26} also included patients who had a short course of radiotherapy (25 Gy) followed by a waiting interval, and three\textsuperscript{18,24,25} included patients who received contact X-ray brachytherapy (60–110 Gy).

In almost half of the studies (11 of 23)\textsuperscript{5,8–10,21,22,24–27,35}, patients were selected for a watchful waiting programme based on the response evaluation after neoadjuvant treatment. In 17 studies\textsuperscript{5,8,9,11–19,22,24,25,27,29–33}, the response evaluation was used to select patients for an additional local excision aiming at organ preservation. Duration of follow-up was reported in 19 studies (83 per cent)\textsuperscript{8,9,11–19,26–29,36}, and varied between a median of 15 and 65.5 months. Treatment characteristics are shown in Table S5.

Response evaluation

The interval between (chemo)radiotherapy and response evaluation varied across studies; most had a 4–12 week interval from the end of (chemo)radiotherapy\textsuperscript{5,8–11,20–23,26–36}. Two studies\textsuperscript{19,25} reported an interval of 14 weeks from the start of neoadjuvant treatment. In two studies\textsuperscript{24,34}, data on timing of the response evaluation was missing. No obvious differences in definitions used for a near-CR were found between studies with a short (less than 8 weeks) versus long (over 8 weeks) interval.

Eighteen studies\textsuperscript{5,8–10,19–23,25–27,28–33,35} defined response based on the combination of DRE and/or endoscopy with radiological imaging. In two studies\textsuperscript{24,36} the response was based on clinical examination alone, and in another three\textsuperscript{11,28,34} only radiological imaging was used for the response assessment. All 21 studies\textsuperscript{5,8–11,19–23,25–35} that included radiological imaging used
MRI to assess the response. In addition to MRI, two studies used endorectal ultrasound imaging and one used FET-CT.

**Terminology**

To describe the response after which patients were considered potentially eligible for watchful waiting or organ preservation, 15 studies used the term ‘near-complete response’. Three studies used the term ‘major response’. Additionally, individual studies used the terms ‘objective clinical response’, ‘suspended complete response’, ‘potential complete response’, ‘subcomplete response’, and ‘small residual lesion’.

**Features indicative of a clinical near-complete response**

In most studies, minor irregularities or a smooth induration found with DRE were considered features of a near-CR, potential complete response, or suspected complete response. Three studies considered the absence of a palpable tumour mass with DRE as a feature of a near-CR or major response. One study, which used DRE as a single modality for response evaluation, considered a decrease in tumour size of more than 50 per cent as a major response.

On endoscopy, 11 studies considered a small flat ulcer or superficial ulceration a feature of a near-CR, potential complete response, major response, or objective clinical response. In four studies, the maximum size of the ulcer (2 or 3 cm) was specified. A small nodule, an irregular mucosa, or small mucosal abnormalities were considered features of a near-CR or suspected complete response in six studies. Another six studies used the term near-CR or small residual lesion when a small residual lesion or tumour with a maximum size of 3 cm was observed.

The response on MRI was based on features observed on a combination of T2W-MRI and DWI in five studies. On T2W-MRI, three studies considered obvious downstaging with or without heterogeneous or irregular residual fibrosis as a feature of a near-CR. Two studies considered a low or intermediate residual signal (not further specified) as feature of a near-CR. None of these studies reported on the extent of downsizing of the tumour. On DWI, four of the five studies considered a small focal area of high signal on high b-value as a feature of a near-CR. The fifth study considered a significant regression of the signal on DWI as feature of a near-CR.

In another nine studies, the response on MRI was based on either the TNM classification (5 studies) or magnetic resonance tumour regression grade (3 studies). In the five studies that used the TNM classification, yCT2 was considered a near-CR or a small residual lesion. In the four studies that used the mrTRG, mrTRG1–2 or mrTRG2 were considered as a near-CR or suspected complete response.

In addition to the luminal response on MRI, nine studies described the response of the lymph nodes. For the response to be classified as a near-CR or major response, or for patients with a small residual lesion to be treated with local excision, the lymph nodes had to be considered not suspicious for residual tumour in seven of nine studies. In two other studies, the response could be classified as a near-CR, if partial regression of the lymph nodes, or obvious downsizing with remaining lymph node(s) of 5 mm or more had occurred. Detailed features that were deemed indicative of a clinical near-CR are provided in Table 5. A summary of findings, including the criteria and features used for each modality, is provided in Table 1.

**Discussion**

This review provides an overview of the terminology, criteria, and features used to describe a near-CR. The terminology used to describe the response category and determine eligibility for watchful waiting or other organ-preserving treatments varied; the most common was ‘near-complete response’, followed by the term ‘major response’. Moreover, criteria and features used to define a near-CR differed between studies. The most common criteria and features for a near-CR were: minor irregularities or a smooth induration found with DRE; small flat ulcer on endoscopy; obvious downstaging of the residual tumour, with or without heterogeneous or irregular residual fibrosis on T2W-MRI; and a small focal area of high signal on DWI. As a consequence of the variation in terminology, criteria, and features, uniformity regarding patient selection for watchful waiting and organ preservation is lacking.

The patients in the included studies were selected for different treatment strategies—either watchful waiting or additional local treatment aiming at organ preservation. The novelty in criteria and features defining a near-CR might be explained by the difference in treatment strategies patients are selected for (watchful waiting or additional local treatment). For selection of patients for watchful waiting, very strict criteria were used initially, which may have missed up to 61 per cent of patients with a pathologic complete response (pCR) at response evaluation before resection. In more contemporary practice, less strict criteria are used to select patients for watchful waiting, with the aim of not missing a cCR in the event of (minor) residual abnormalities at response evaluation. These patients still have a high likelihood of achieving a cCR. In contrast, to select patients for additional local treatment aiming at organ preservation, the probability of the response evolving into a cCR is less important, and the criteria used to select patients are more liberal than those used to select patients for watchful waiting. Instead, patients are selected based on the risk of recurrence after additional local treatment, associated with specific high-risk histopathological features and tumour stage.

The findings of this review might indicate that there is a need for a new response categorization system accompanied by new terminology that can differentiate the response in a way that allows clinicians to distinguish patients who are candidates for either watchful waiting or additional local treatment aiming at organ preservation. The term ‘near-complete response’ should perhaps be reserved for patients with a degree of response with high potential for evolving into a cCR. A different term can then be used to describe a major response but with persistent abnormalities that are likely indicative of minor residual tumour requiring additional treatment. An alternative approach could be to convert verbal terminology into numbered response categories, analogous to radiology scoring systems such as Breast Imaging - Reporting and Data System (BI-RADS) and Prostate Imaging - Reporting and Data System (PI-RADS). In such a system, the different response categories should correspond to the likelihood of a sustained complete response or the likelihood of regrowth.

An underlying issue with defining a near-CR is the limited evidence for the predictive value of a cCR or residual tumour for both individual features as well as a combination of features from different modalities. Research groups interested in organ preservation will, therefore, develop their own definition based on clinical experience, which will inherently lead to
heterogeneity and bias. The results of the present review indicated that, on endoscopy, a small flat ulcer was considered a sign indicative of a near-CR in most studies. In a study by van der Sande et al., the positive predictive value for a true and persistent complete response of a small ulcer was 40–50 per cent. In addition, the predictive value of a flat white scar was 70–80 per cent and that of a large ulcer was 29–30 per cent. It is important to note that, when a large ulcer is observed, the risk of residual disease is substantially higher than the chance of a cCR.

In this review, features commonly used to describe a near-CR on MRI were obvious downstaging with or without heterogeneous or irregular fibrosis on T2W-MRI and small focal spots of high signal on DWI. Lambrechts et al. assessed four response patterns on T2W-MRI and DWI, for which positive predictive values for a cCR were 0, 88, 89, and 100 per cent. The two response patterns with a positive predictive value of 88 and 89 per cent, semicircular fibrosis and regression of polyloid tumours with fibrosis at the stalk (both regardless of any high diffusion signal), might be considered a near-CR. However, using standalone modalities (either endoscopy or MRI) will inherently lead to underestimation or overestimation of the tumour response. Therefore, it is very important to perform a multimodal assessment using DRE, endoscopy, and MRI.

Lymph node status is also important in defining a near-CR. In the present review, less than half of the studies using radiological imaging described the criteria for a lymph node response. This is remarkable, because the locoregional lymph nodes remain in situ in organ-preserving treatments, and persisting lymph node metastases will increase the risk of recurrence. Even though it is more accurate than primary staging, lymph node staging after (chemo)radiotherapy remains challenging. Although not adopted universally, the main criterion used for lymph node staging is the short-axis size of the lymph nodes, for which a 5-mm cut-off has been recommended. Although this cut-off provides reasonable results, overstaging and understaging after (chemo)radiotherapy are still encountered. As the lymph nodes are small, their morphology is more difficult to evaluate. This results in a sensitivity of around 40 per cent for the detection of lymph node metastases in patients with a good tumour response after (chemo)radiotherapy. The risk of lymph node metastases after (chemo)radiotherapy should not be underestimated, as a study by Haak et al. reported that the risk of positive lymph nodes was 7 per cent for ypT0 and 12 per cent for ypT1 tumours. However, as surveillance of suspicious lymph nodes can be done easily with MRI, and salvage treatment is rarely jeopardized by lymph node metastasis, the authors encourage offering an organ-preserving treatment strategy even to patients with lymph nodes of at least 5 mm and a luminal (near-)CR, provided that the lymph nodes are monitored on careful imaging during follow-up. Lymph node status should, therefore, be integrated into the definition of a near-CR.

Owing to the heterogeneity of criteria and features identified, a clear uniform definition of a near-CR cannot be proposed based on the present review. The most important next step is to gather evidence regarding which features at response evaluation are highly predictive of successful watchful waiting or successful organ preservation after additional local treatment. However, pending such evidence, a first step could be to establish expert-based consensus. This may take into account two potential response categories: the response with a high likelihood of progression into a cCR, after which watchful waiting can be considered; and a major response with the possibility of residual tumour, after which additional local treatment aiming at organ preservation can be considered.

There are some limitations to this work. The studies included in this review are heterogeneous with regard to baseline characteristics of the patients, neoadjuvant treatment schedules, and the timing of response evaluation, complicating comparisons. In addition, multiple outcome measures were used in the studies. Therefore, features and criteria used to describe a near-CR could not be compared by clinical outcome. Some included studies are from the same research group, with the potential for reporting on overlapping cohorts. This might have led to over-representation of some features in this review. Consideration was given to excluding studies from the same research group or studies with potentially overlapping cohorts. However, this may have missed important data because the terminology and/or definition used by a research group may have changed over time, owing to changes in management and advances in evidence and knowledge.

Funding
The authors have no funding to declare.

Acknowledgements
This review was not registered and no protocol was prepared.

Disclosure
The authors declare no conflict of interest.

Supplementary material
Supplementary material is available at BJS online.

Data availability
No new data were generated or analysed in this manuscript.

References


