



**Universiteit  
Leiden**  
The Netherlands

## **Predictive value of endoscopic features for a complete response after chemoradiotherapy for rectal cancer**

Sande, M.E. van der; Maas, M.; Melenhorst, J.; Breukink, S.O.; Leerdam, M.E. van; Beets, G.L.

### **Citation**

Sande, M. E. van der, Maas, M., Melenhorst, J., Breukink, S. O., Leerdam, M. E. van, & Beets, G. L. (2021). Predictive value of endoscopic features for a complete response after chemoradiotherapy for rectal cancer. *Annals Of Surgery*, 274(6), E541-E547.  
doi:10.1097/SLA.0000000000003718

Version: Publisher's Version  
License: [Creative Commons CC BY-NC-ND 4.0 license](#)  
Downloaded from: <https://hdl.handle.net/1887/3484022>

**Note:** To cite this publication please use the final published version (if applicable).

# Predictive Value of Endoscopic Features for a Complete Response After Chemoradiotherapy for Rectal Cancer

Marit E. van der Sande, MD,\*✉ Monique Maas, MD, PhD,† Jarno Melenhorst, MD, PhD,‡  
Stéphanie O. Breukink, MD, PhD,‡ Monique E. van Leerdam, MD, PhD,§ and Geerard L. Beets, MD, PhD\*✉

**Objective and Background:** Watch-and-wait approach in rectal cancer relies on the identification of a clinical complete response (CR) after neoadjuvant (chemo)radiotherapy. This is mainly performed by rectal examination, magnetic resonance imaging, and endoscopy. Endoscopy has been less well studied, and the objective of the study is to assess the diagnostic value of endoscopy and the predictive value of endoscopic features for the identification of CR.

**Patients and Methods:** A total of 161 patients with primary rectal cancer undergoing flexible sigmoidoscopy for response assessment after neoadjuvant (chemo)radiotherapy between January 2012 and December 2015 at a single institution were evaluated retrospectively. Three independent readers scored endoscopic features and a confidence level score for a CR. Diagnostic performance of endoscopy and positive predictive value (PPV) of endoscopic features for a CR were calculated. If available, biopsy results were revealed to the reader and a change in confidence level was noted. Reference standard was histology after surgery, or long-term outcome in a watch-and-wait policy.

**Results:** Median time to endoscopy was 9 (interquartile range 8–12) weeks. Area under the receiver operator characteristic curve, sensitivity, specificity, PPV, and negative predictive value for a CR were 0.80 to 0.84, 72% to 94%, 61% to 85%, 63% to 78% and 80% to 89%, respectively. A flat scar was the most predictive feature of a CR (PPV 70%–80%). The PPV of small flat ulcers and large flat ulcers were 40% to 50% and 29% to 33%, respectively. The addition of biopsy results led to a significant change in confidence level score in 4% to 13% of patients.

**Conclusions:** More than 70% of the patients with a luminal CR after neoadjuvant treatment for rectal cancer can be identified by endoscopy at  $\pm 9$  weeks. Together with findings on digital rectal examination (DRE) and magnetic resonance imaging, specific endoscopic features can be used to select patients for an extended observation period to select for organ preservation.

**Keywords:** endoscopy, organ preservation, rectal cancer, response assessment

(Ann Surg 2021;274:e541–e547)

The introduction of the total mesorectal excision (TME), preoperative radiotherapy, and chemoradiotherapy (CRT) in the treatment of rectal cancer have led to a marked improvement in the local recurrence rates.<sup>1–3</sup> The current focus in rectal cancer management

has therefore shifted from improving oncological outcomes toward reducing the surgical morbidity and the adverse effects on functional outcome. A watch-and-wait approach, that is, no immediate surgery and close surveillance, could potentially avoid the risk for postoperative complications and significantly benefit quality of life in patients with a complete response (CR) to neoadjuvant treatment.<sup>4,5</sup>

Although recent publications have shown promising outcomes with a low oncological risk,<sup>6,7</sup> a few challenges remain to be overcome before widespread implementation of the watch-and-wait approach for rectal cancer can take place in clinical practice. One of these challenges is improving the selection of patients with rectal cancer who are candidates for a watch-and-wait approach, for which accurate methods for tumor response assessment after neoadjuvant treatment are needed. Recommended modalities for tumor response assessment include imaging with magnetic resonance imaging (MRI) and clinical assessment with digital rectal examination, endoscopy, and biopsies.<sup>6,8</sup> Response assessment with MRI is performed by T2-weighted and diffusion weighted imaging, and this combination yields a reasonable diagnostic accuracy for detection of a CR. The overall endoscopic and clinical assessment has been shown to be more accurate than MRI in 1 study, albeit with the same limitation of a relatively low sensitivity for the detection of a CR.<sup>9</sup> One of the reasons for this is that patients with a pathological CR may show residual ulcers or irregularities at restaging after the recommended 8- to 12-week interval from neoadjuvant treatment, instead of the white scar which is typically associated with a clinical CR.<sup>10,11</sup> In these cases, MRI will usually show ambiguous abnormalities (eg, heterogeneous fibrosis or small focal diffusion restriction), which does not really help in clinical decision making. To solve this problem, the observation interval is increasingly extended with a second restaging after another 6 to 12 weeks in patients with a near clinical CR, to allow for further regression.<sup>12</sup> It is, however, unclear which group of patients will benefit from such an extended interval. Possibly, distinct features on endoscopy can be used to define groups of patients that may benefit from an extended interval or inclusion of a watch-and-wait policy. Therefore, the aims of this study were to assess the overall diagnostic accuracy of endoscopy for response assessment and to assess the predictive value for a CR of various endoscopic features.

## MATERIALS AND METHODS

### Study Population

This study was performed in a single center in which restaging endoscopy was routinely done after neoadjuvant treatment for rectal cancer to assess the luminal tumor response.

All consecutive patients with primary rectal cancer who received neoadjuvant treatment between January 2012 and December 2015 were identified. Inclusion criteria were (1) primary rectal cancer; (2) neoadjuvant treatment with short-course radiotherapy or long-course CRT, both followed by a waiting interval to allow for downsizing; (3) restaging endoscopy within 6 to 16 weeks after the

From the \*Department of Surgery, The Netherlands Cancer Institute, Amsterdam, The Netherlands; †Department of Radiology, The Netherlands Cancer Institute, Amsterdam, The Netherlands; ‡Department of Surgery, Maastricht University Medical Centre, Maastricht, The Netherlands; and §Department of Gastroenterology, The Netherlands Cancer Institute, Amsterdam, the Netherlands.

✉ g.beets@nki.nl; m.vd.sande@nki.nl.

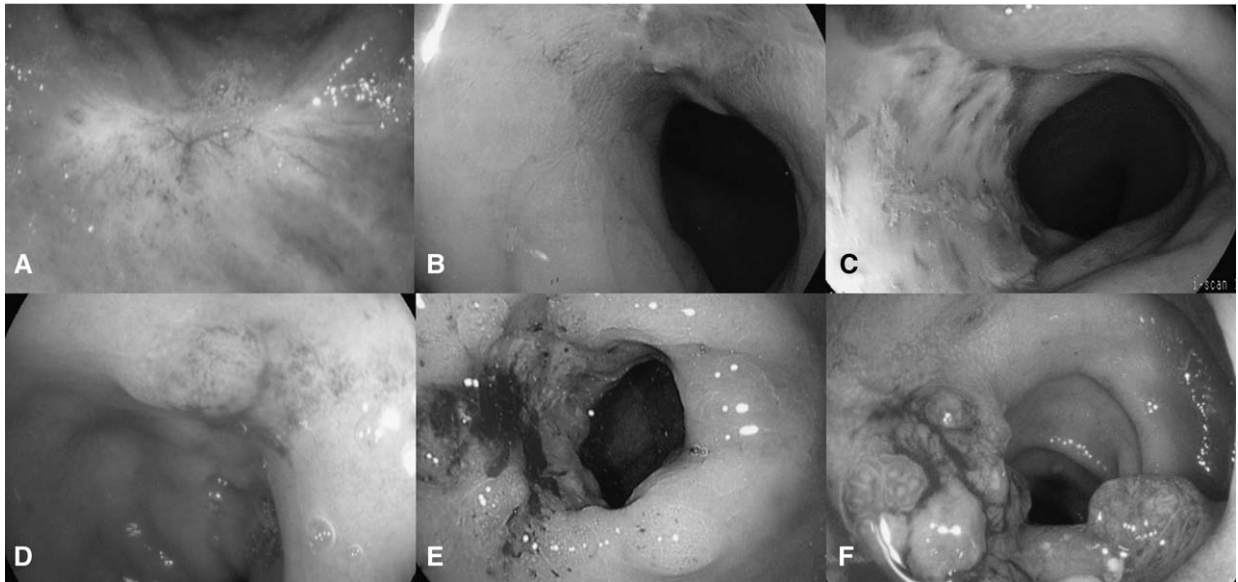
The authors report no conflicts of interest.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.annalsofsurgery.com).

Copyright © 2019 Wolters Kluwer Health, Inc. All rights reserved.

ISSN: 0003-4932/19/27406-e541

DOI: 10.1097/SLA.0000000000003718



**FIGURE 1.** Endoscopic features at response assessment. A, Flat scar, (B) small flat ulcer <1 cm, (C) large flat ulcer, (D) adenomatous mass, (E) ulcer with an irregular border, (F) tumorous mass.

end of neoadjuvant treatment; and (4) interval of less than 8 weeks between endoscopy and surgery. Exclusion criteria consisted of (1) unresectable rectal cancer, (2) absence of stored endoscopy images, and (3) insufficient endoscopy image quality (eg, view impaired by residual stool, poor photo resolution, unclear location of remnant). After restaging, patients either underwent rectal cancer surgery or were enrolled in a watch-and-wait program as part of a prospective cohort study (registered on clinicaltrials.gov: NCT00939666 and NCT02278653). Follow-up data of more than 2 years were available for all patients undergoing a watch-and-wait approach. This study was approved by the Committee on Medical Research Ethics and informed consent was waived due to the retrospective nature of this study.

### Endoscopic Assessment

All patients underwent flexible sigmoidoscopy (EPK-i video processor, Pentax Medical Netherlands, Uithoorn, The Netherlands) 6 to 16 weeks after the completion of neoadjuvant treatment for luminal response assessment. Pre-examination bowel cleansing was done with a rectal phosphate enema. Only white light imaging was used and images were digitally stored. Biopsies during endoscopy were not standard clinical routine, but were taken only at the endoscopists' judgment.

### Endoscopic Image Evaluation

All restaging endoscopic images were retrospectively and independently evaluated by 3 readers. R1 (G.L.B.) is a gastrointestinal surgeon with 13 years of experience in assessing restaging endoscopy. R2 (J.M.) is a gastrointestinal surgeon with 4 years of experience in assessing restaging endoscopy. R3 (M.E.L.) is a gastroenterologist with 3 years of experience in assessing restaging endoscopy and 14 years of experience in diagnostic colonoscopy. Colonoscopy images of the primary tumor before neoadjuvant treatment were not at the readers' disposal. The readers were blinded to each other's results and to the patients' outcomes.

The readers were asked to score the presence of the following endoscopic features, and to assign the most predominant endoscopic feature of each patient's endoscopic images:

1. a flat scar
2. a small flat ulcer (<1 cm)
3. a large flat ulcer
4. ulcer with an irregular border
5. an adenomatous mass
6. tumorous mass.

The endoscopic features are presented in Figure 1.

The readers were also asked to score the likelihood of a CR based on the endoscopic images, using a confidence level score (1=definitely residual tumor, 2=probably residual tumor, 3=possibly residual tumor/possibly CR, 4=probably CR, 5=definitely CR). In case biopsies were taken at restaging endoscopy, the results were shared with the readers after their first visual evaluation and they were asked again to give a confidence level for a CR based on the images and biopsy results together. A change in confidence level score based on biopsy results was noted.

### Reference Standard

The histopathological staging of the surgical resection specimen served as the reference standard in patients undergoing surgery. The surgical specimens were assessed according to the method of Quirke and Dixon,<sup>13</sup> and response was described using the TNM staging classification. Because this study focused on the luminal response assessment, N (nodal) stage was not included and only T (tumor) stage was used as a reference standard for a CR. A ypT0 was considered a CR, a ypT1–4 was considered residual tumor. In patients undergoing a watch-and-wait policy, a sustained CR of more than 2 years, without evidence of luminal regrowth, was considered a CR. This period for a sustained CR was chosen as the vast majority of local regrowth appear within the first 2 years of follow-up and are rarely encountered after this period.<sup>6,14</sup>

### Statistical Analyses

Statistical analyses were performed using SPSS Statistics version 22.0 (IBM Corporation, Armonk, NY) and STATA version 11 (StataCorp LLC, TX). Nominal data are presented as absolute frequencies and percentages. Continuous data are presented as

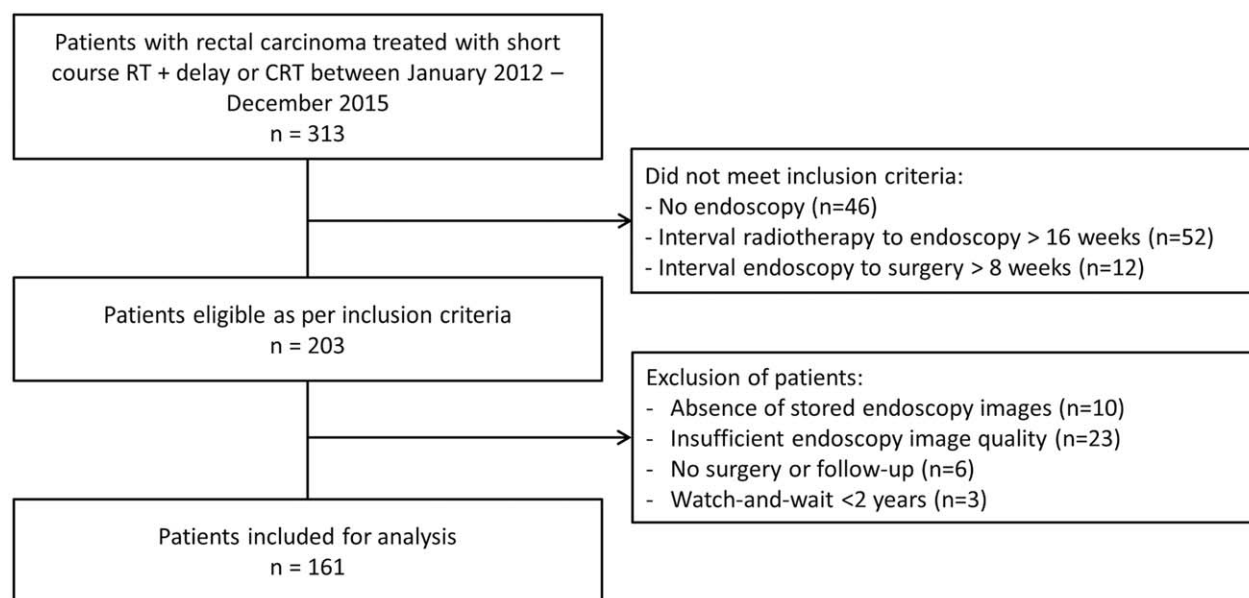


FIGURE 2. Flowchart of patient inclusion.

median numbers with interquartile range (IQR). Receiver operating characteristic (ROC) curves were plotted for each reader to assess the diagnostic value of endoscopy for the detection of a CR. Cut-off for the confidence was set between confidence levels 3 and 4, which was decided before the start of the study. CR was the positive outcome measure. The area under the ROC curve (AUC), sensitivity, specificity, positive predictive value (PPV), and negative predictive value with 95% confidence intervals (CIs) were calculated for each reader. PPVs for a CR were calculated for each endoscopic feature. AUCs were compared by the method of Hanley and McNeil.<sup>15</sup> Quadratic weighted kappa was calculated to assess the interobserver agreement between readers for the confidence level of a CR. Cohen kappa was calculated to assess the interobserver agreement for the categorization of endoscopic features (0–0.20 poor; 0.21–0.40 fair; 0.41–0.60 moderate; 0.61–0.80 substantial; 0.81–1 excellent agreement).

## RESULTS

### Patients

Three hundred thirteen patients with rectal cancer treated with CRT or short-course radiotherapy with a waiting interval between January 2012 and December 2015 were identified, of whom 203 met the inclusion criteria. Figure 2 presents the patient flowchart. Of these 203 patients, we excluded 42 patients for the following reasons: absence of stored endoscopy images ( $n = 10$ ), insufficient endoscopic image quality ( $n = 23$ ), no surgery or follow-up ( $n = 6$ ), and follow-up in watch-and-wait less than 2 years ( $n = 3$ ). A total of 161 patients, of whom 104 (65%) men, were included in the analysis. Median age was 65 (IQR 57–73) years. Baseline patient characteristics are shown in Table 1.

Sigmoidoscopy was performed after a median 9 (IQR 8–12) weeks after the last day of neoadjuvant radiotherapy. Eighty-seven (54%) patients subsequently underwent rectal surgery, of whom 80 patients underwent TME and 7 patients underwent local excision with transanal endoscopic microsurgery. Fourteen (16%) of the 87 patients undergoing rectal surgery had ypT0 and 73 (84%) patients had residual luminal tumor (ypT1–4) at pathologic assessment. One out of 9 patients with luminal CR after TME had a pathological

positive lymph node (ypT0N1), but was considered a luminal CR for this study. Seventy-four (47%) patients underwent a watch-and-wait policy, of which 19 (26%) experienced a luminal regrowth after a median follow-up time of 6 (IQR 4–18) months. Fifty-five (74%) patients remained without luminal regrowth, with a median follow-up time of 45 (IQR 34–53) months. In the patients undergoing rectal surgery, surgery was performed a median of 18 (IQR 15–35) days after the sigmoidoscopy.

### Diagnostic Accuracy of Endoscopy

Endoscopy was accurate in predicting the response in 126 (78%), 118 (73%), and 129 (80%) of 161 patients for R1, R2, and R3, respectively. The ROC curves for the prediction of CR of all 3 readers are presented in Figure 3. The AUCs for the prediction of CR were 0.84 (95% CI: 0.77–0.90), 0.80 (95% CI: 0.73–0.87), and 0.84 (95% CI: 0.78–0.91) for R1, R2, and R3, respectively. Table 2 shows all diagnostic accuracy parameters.

### Endoscopic Features

The distribution of endoscopic features was significantly different between patients with and without a CR for each reader ( $P < 0.001$ ). In patients with a CR, 64% to 77% had a flat scar, 6% to 12% a small flat ulcer, 6% to 10% a large flat ulcer, 1% to 7% an ulcer with an irregular border, 4% an adenomatous mass, and 4% to 6% gross residual mass. In contrast, patients without CR showed a flat scar in 12% to 25%, small flat ulcer in 7% to 11%, a large flat ulcer in 11% to 15%, an ulcer with an irregular border in 19% to 30%, an adenomatous mass in 7% to 20%, and gross residual mass in 19% to 28%.

Table 3 provides the PPV for a CR per endoscopic feature. A flat scar was most predictive of a CR: 70% to 80% of patients had a CR. Small flat ulcers and large flat ulcers were predictive of a CR in 40% to 50% and 29% to 33%, respectively.

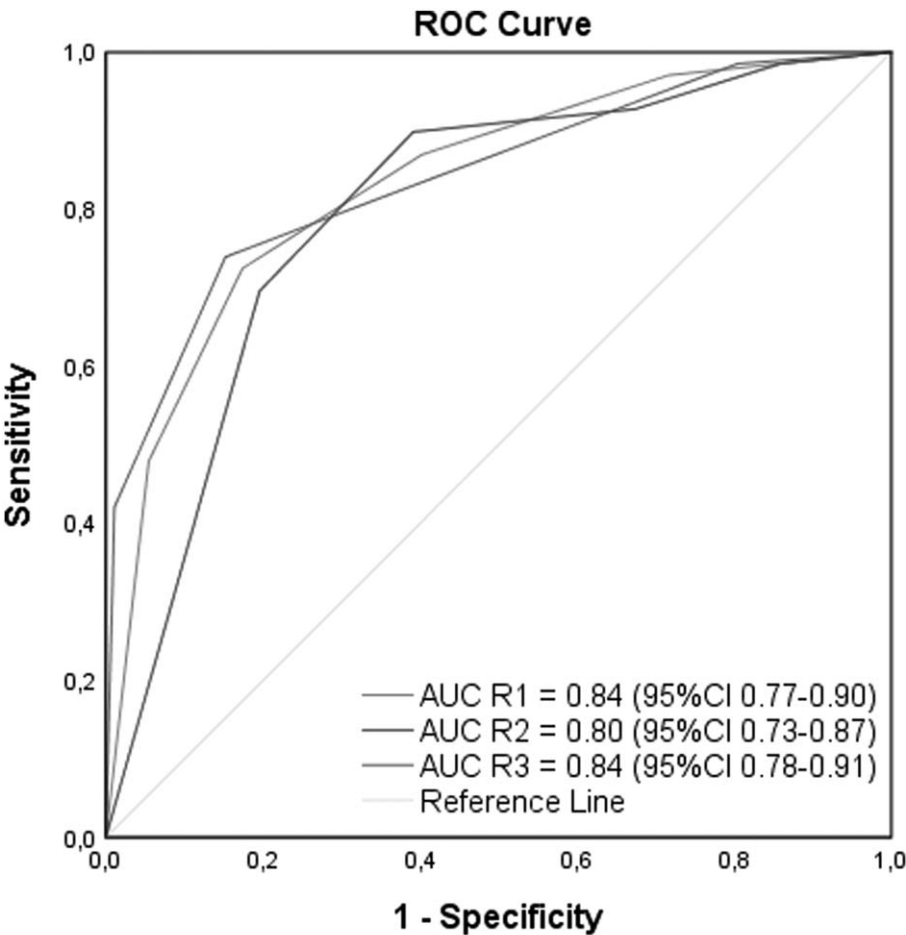
### Biopsies

Biopsies were taken in 55 of 161 patients (34%). Biopsy histology showed no (pre)malignant features in 34 (62%) patients, low-grade dysplasia in 6 (11%) patients, high-grade dysplasia in

TABLE 1. Patient With Rectal Cancer and Treatment Characteristics

|                                                    | All (n = 161) | Non-CR (n = 92) | CR (n = 69) | P      |
|----------------------------------------------------|---------------|-----------------|-------------|--------|
| Sex, n (%)                                         |               |                 |             | 0.849  |
| Male                                               | 104 (64.6)    | 60 (65.2)       | 44 (63.8)   |        |
| Female                                             | 57 (35.4)     | 32 (34.8)       | 25 (36.2)   |        |
| Age, median (IQR), yr                              | 64 (57–73)    | 64 (57–73)      | 64 (57–72)  |        |
| Clinical T-stage, n (%)                            |               |                 |             | 0.337  |
| cT1–2                                              | 35 (21.7)     | 17 (18.5)       | 18 (26.1)   |        |
| cT3                                                | 118 (73.3)    | 69 (75.0)       | 49 (71.0)   |        |
| cT4                                                | 8 (5.0)       | 6 (6.5)         | 2 (2.9)     |        |
| Clinical N-stage, n (%)                            |               |                 |             | 0.027  |
| cN0                                                | 41 (25.5)     | 24 (26.1)       | 17 (24.6)   |        |
| cN1                                                | 46 (28.6)     | 19 (20.7)       | 27 (39.1)   |        |
| cN2                                                | 74 (46.0)     | 49 (53.3)       | 25 (36.2)   |        |
| Neoadjuvant treatment, n (%)                       |               |                 |             | 0.001  |
| Long course CRT                                    | 145 (90.1)    | 77 (83.7)       | 68 (98.6)   |        |
| Short course RT + interval                         | 16 (9.9)      | 15 (16.3)       | 1 (1.4)     |        |
| Type of treatment, n (%)                           |               |                 |             | <0.001 |
| Surgery                                            | 87 (54.0)     | 73 (79.3)       | 14 (20.3)   |        |
| Watch-and-wait                                     | 74 (46.0)     | 19 (20.7)       | 55 (79.7)   |        |
| Interval RT to endoscopy, median (IQR), weeks      | 9 (8–12)      | 8 (8–10)        | 10 (8–13)   | 0.001  |
| Interval endoscopy to surgery, median (IQR), days* | 18 (15–35)    | 18 (13–30)      | 34 (15–44)  | 0.055  |

\*In patients undergoing surgery.  
RT indicates radiotherapy.



Diagonal segments are produced by ties.

FIGURE 3. AUC for the prediction of complete response (n = 161).

**TABLE 2.** Diagnostic Accuracy of Endoscopy Using a Confidence Level Score for the Prediction of a Complete Response

|             | Reader 1   | Reader 2   | Reader 3   |
|-------------|------------|------------|------------|
| Sensitivity | 72 (50/69) | 90 (62/69) | 74 (51/69) |
| 95% CI      | 62–83      | 83–97      | 64–84      |
| Specificity | 83 (76/92) | 61 (56/92) | 85 (78/92) |
| 95% CI      | 75–90      | 51–71      | 77–92      |
| PPV         | 76 (50/66) | 63 (62/98) | 78 (51/65) |
| 95% CI      | 65–86      | 54–73      | 68–88      |
| NPV         | 80 (76/95) | 89 (56/63) | 81 (78/96) |
| 95% CI      | 72–88      | 81–97      | 73–89      |
| AUC         | 0.84       | 0.80       | 0.84       |
| 95% CI      | 0.77–0.90  | 0.73–0.87  | 0.78–0.91  |

Numbers are percentages. Absolute numbers are given within parentheses.  
CI indicates confidence interval; NPV, negative predictive value.

7 (13%) patients, and adenocarcinoma in 8 (15%) patients. Thirty out of 55 patients with a biopsy had residual tumor after resection or a watch-and-wait policy, of whom 7 (23%) patients had a biopsy positive for adenocarcinoma.

After the biopsy results were revealed to the readers, the confidence level for a CR changed in 15 (27%), 18 (33%), and 13 (24%) out of 55 patients for R1, R2, and R3, respectively. In 4 (7%), 7 (13%), and 2 (4%) patients the confidence level changed from residual tumor to CR or vice versa (for details see Supplemental Digital Content Table 1, <http://links.lww.com/SLA/B844>). The AUCs did not significantly change for the 3 readers for sigmoidoscopy with biopsy results compared to sigmoidoscopy without biopsy results in these 55 patients (data shown in Supplemental Digital Content Fig. 1, <http://links.lww.com/SLA/B844>).

### Interobserver Agreement

Table 4 shows the interobserver agreement between the 3 readers for both confidence level and endoscopic features. The readers showed substantial agreement in the confidence level for a CR ( $\kappa$  0.62–0.76). Lower agreement was shown when the readers

**TABLE 4.** Interobserver Agreement of 3 Readers on Endoscopy for the Assignment of a Confidence Level or Endoscopic Feature

|                   | Reader 1–Reader 2 | Reader 1–Reader 3 | Reader 2–Reader 3 |
|-------------------|-------------------|-------------------|-------------------|
| Confidence Level* | 0.67 (0.59–0.76)  | 0.76 (0.70–0.83)  | 0.62 (0.52–0.72)  |
| Features†         | 0.54 (0.45–0.62)  | 0.64 (0.55–0.72)  | 0.47 (0.39–0.56)  |

\*Kappa values with quadratic weighting.

†Cohen kappa values.

assigned the individual endoscopic features, with moderate to substantial agreement ( $\kappa$  0.47–0.64).

## DISCUSSION

In the present study the overall judgment of the endoscopy reader of a definite or probable CR identified 72% to 90% of the patients with a luminal CR (sensitivity). Of those assessed as definite or probable complete responders, 63% to 78% truly had a pathological CR or were free of regrowth in a watch-and-wait approach (PPV). The present study also showed that although the majority of patients with a CR presented with the typical flat scar, 23% to 36% showed other endoscopic features at response assessment. The PPV for a CR was highest in patients with a flat scar (70%–80%), and was 40% to 50% in patients with a small flat ulcer. Large flat ulcers, ulcers with an irregular border, or residual adenomatous or tumorous masses were more consistent with residual tumor, as we found lower predictive values for a CR in these patients.

There are relatively few studies that have assessed the accuracy of endoscopic findings to identify pCR. One early study investigating the utility of proctoscopy for response assessment reported disappointing results with poor accuracy, identifying only 50% of complete responders and with 25% residual tumor in those thought to be CR.<sup>16</sup> In another study, flattening of the marginal tumor swelling and re-epithelialization of ulceration was assessed on endoscopic images to predict pCR.<sup>17</sup> The sensitivity of these findings to predict a pCR was 69% to 87%, but specificity was only 39% to 74%. In both studies response assessment was, however, performed within 6 weeks after CRT. In a 2015 article by our group, clinical assessment with endoscopy and DRE was more accurate than MRI alone for identification of a CR, with a sensitivity of 53% and specificity of 97%.<sup>9</sup> Residual tumor was rarely missed with endoscopy and DRE, at the expense of a lower sensitivity to detect CR compared to the present study. This demonstrates that the readers in the present study have adopted a selection strategy aimed at identifying more responders, accepting a higher rate of false positive CRs (22%–37%), a change of strategy that is discussed below.

Habr-Gama et al,<sup>11</sup> the pioneers of watch-and-wait, have previously described the frequently observed endoscopic features in complete responders: whitening of the mucosa with telangiectasia and absence of any mucosal, superficial or deep ulceration, any palpable nodule, or stenosis. Many organ preservation trials have adopted these criteria for a clinical CR. In a recent study only 27% of the patients with a pCR, however, fulfilled these criteria at preoperative restaging.<sup>18</sup> Accordingly, previous studies have shown that when the mucosa was assessed after resection, only 25% to 39% of patients with pCR showed mucosal integrity or white scarring.<sup>19–21</sup> In our study, 64% to 77% of patients with CR had a flat scar as the predominant endoscopic feature, whereas 23% to 36% of the complete responders had other endoscopic presentations. The original

**TABLE 3.** Predictive Value for a Complete Response of Endoscopic Features

|                     | Reader 1            | Reader 2            | Reader 3            |
|---------------------|---------------------|---------------------|---------------------|
| Flat scar           |                     |                     |                     |
| PPV for CR (95% CI) | 77 (44/57)<br>66–88 | 70 (53/76)<br>59–80 | 80 (45/56)<br>70–91 |
| Small flat ulcer    |                     |                     |                     |
| PPV for CR (95% CI) | 50 (7/14)<br>24–76  | 40 (4/10)<br>9–70   | 44 (8/18)<br>21–67  |
| Large flat ulcer    |                     |                     |                     |
| PPV for CR (95% CI) | 33 (5/15)<br>9–57   | 29 (4/14)<br>5–52   | 33 (7/21)<br>13–54  |
| Irregular ulcer     |                     |                     |                     |
| PPV for CR (95% CI) | 15 (5/33)<br>3–27   | 15 (3/20)<br>0–31   | 4 (1/27)<br>0–11    |
| Adenoma             |                     |                     |                     |
| PPV for CR (95% CI) | 27 (3/11)<br>1–54   | 14 (3/21)<br>0–29   | 33 (3/9)<br>3–64    |
| Tumor mass          |                     |                     |                     |
| PPV for CR (95% CI) | 13 (4/30)<br>1–26   | 11 (2/19)<br>0–24   | 14 (4/29)<br>1–26   |

Numbers are percentages. Absolute numbers are given within parentheses.  
CI indicates confidence interval.

Habr-Gama criteria for a clinical CR represent a very strict approach for the selection of patients for a watch-and-wait policy, yet evidently underestimate the tumor response. When exclusively patients with a flat scar without any mucosal abnormalities are considered for organ preservation, many patients with a good (and possibly complete) response will be managed by radical surgery, losing the opportunity for organ preservation. Widening the inclusion criteria for a watch-and-wait policy, for example, including patients with a residual ulcer, will lead to identifying more patients with a CR. Of course, this will also lead to an increased number of patients with residual tumor or local regrowth during follow-up. This strategy can also be interpreted as extending the observation interval to give tumors with a very good response the opportunity to regress further, and to decide later whether or not to proceed with a resection.<sup>12</sup> In the present study half of the patients with a small flat ulcer had either a pCR after resection or a sustained CR during watch-and-wait, and an extended observation interval could have identified these patients. On the contrary, in patients who showed a residual large ulcer, ulcer with an irregular border, adenomatous mass, or gross tumorous mass at endoscopy lower predictive values for a CR were found and direct surgery seems justified.

Although the present study evaluated response assessment after neoadjuvant treatment with endoscopy only, in clinical practice the response assessment and the decision-making process also relies on important additional information of the digital rectal examination and MRI.<sup>6,9</sup> MRI provides information on both the response of the primary tumor (with morphologic patterns and MRI tumor regression grading) and on nodal response.<sup>22</sup> Recent guidelines recommend MRI to be routinely performed in every patient after CRT.<sup>8,23</sup> Additional considerations in decision making are the distance of the tumor to the sphincter, and how the patient values sphincter and organ preservation. The present study was not focused on this clinical decision-making process, but rather on the value of the different features of the endoscopic response. It is, however, clear that this information should be combined with MRI findings, as 1 study showed that this combination yielded the highest accuracy with an AUC under the ROC curve of 0.89.<sup>9</sup> Further research should define the chances of a CR for the different combinations of findings on different assessment techniques.

The interobserver agreement for the presence of the endoscopic features and the confidence level scores for a CR were moderate to substantial between the 3 readers. Two of the readers (R1 and R3) work at the same hospital and perform restaging with continuous feedback to each other and had substantial agreement in confidence levels for CR prediction. The agreement between R2 and the 2 other readers was lower, because R2 showed a more liberal approach with higher confidence levels. This resulted in a sensitivity of more than 90%, but with a higher false positive rate of 37%.

The added value of biopsies is questionable. Interestingly, in our study the addition of biopsies almost never changed the readers' confidence level of response. The AUC for endoscopy with biopsies was not statistically different from the AUC of endoscopy alone. This may have been caused by a lack of confidence in biopsy results from the readers, as previous reports have shown that negative biopsies often represent sampling errors and are therefore of limited clinical use for ruling out residual tumor.<sup>24,25</sup> This is supported by our results that showed false negative biopsy results in 50% of the patients.

Advanced endoscopy technologies such as the narrow-spectrum technologies and autofluorescence imaging are an expanding field of research and may improve the mucosal, structural, and microvascular visualization of the rectal wall.<sup>26</sup> This may help in the differentiation between ongoing response or residual tumor after neoadjuvant treatment in rectal cancer. These technologies have, however, yet to be explored in the setting of restaging assessment.

Several limitations of this study should be addressed. This was a retrospective study and the readers only evaluated the static endoscopic images taken by other endoscopists; video was not available. We therefore excluded all patients in whom the readers felt the images were of insufficient quality to make a prediction, potentially leading to selection bias. Also, this study was performed in a referral center for organ preservation. Patients with a good response mostly established on MRI were referred to our center for full response assessment, leading to a high a priori chance of a CR in this study. This has an influence on the positive and negative predictive values. In the current study, nodal involvement was disregarded as endoscopy only assesses the luminal response and not nodal response, whereas in clinical practice nodal involvement and other MRI findings will obviously have influence on the treatment. Last, the range of the intervals between neoadjuvant treatment and endoscopy, and between endoscopy and surgery might also have influenced our results, although the AUCs were not significantly different between patients with an interval shorter or longer than the median interval to endoscopy and to surgery, as shown in Supplementary Digital Content Tables 2 and 3, <http://links.lww.com/SLA/B844>.

In conclusion, our study demonstrates that more than 70% of the patients with a luminal CR after neoadjuvant treatment for rectal cancer can be identified by endoscopy. Although the majority of patients with a CR presented with the typical flat scar, one third showed other endoscopic features at response assessment. When a small flat ulcer was seen 40% to 50% went on to have a CR, and these patients may be selected for an extended observation interval. For patients with large flat ulcers, ulcers with an irregular border or residual adenomatous or tumorous masses, predictive values for a CR were low and direct surgery is justified. When combined with the findings of clinical examination and restaging MRI, this information can be of value when discussing the treatment option of organ preservation with patients who have a good response to chemoradiation.

## REFERENCES

1. Heald RJ, Ryall RD. Recurrence and survival after total mesorectal excision for rectal cancer. *Lancet*. 1986;1:1479–1482.
2. Swedish Rectal Cancer Trial, Cedermark B, Dahlberg M, Glimelius B, et al. Improved survival with preoperative radiotherapy in resectable rectal cancer. *N Engl J Med*. 1997;336:980–987.
3. Van Gijn W, Marijnen CA, Nagtegaal ID, et al. Preoperative radiotherapy combined with total mesorectal excision for resectable rectal cancer: 12-year follow-up of the multicentre, randomised controlled TME trial. *Lancet Oncol*. 2011;12:575–582.
4. Habr-Gama A, Perez RO, Nadalin W, et al. Operative versus nonoperative treatment for stage 0 distal rectal cancer following chemoradiation therapy: long-term results. *Ann Surg*. 2004;240:711–717. discussion 717–718.
5. Hupkens BJP, Martens MH, Stoot JH, 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*. 2017;60:1032–1040.
6. Van der Valk MJM, Hilling DE, Bastiaannet E, et al. Long-term outcomes of clinical complete responders after neoadjuvant treatment for rectal cancer in the International Watch & Wait Database (IWWD): an international multicentre registry study. *Lancet*. 2018;391:2537–2545.
7. Dattani M, Heald RJ, Goussous G, et al. Oncological and survival outcomes in watch and wait patients with a clinical complete response after neoadjuvant chemoradiotherapy for rectal cancer: a systematic review and pooled analysis. *Ann Surg*. 2018;268:955–967.
8. Glynne-Jones R, Wyrwicz L, Tiret E, et al. Rectal cancer: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. *Ann Oncol*. 2017;28:iv22–iv40.
9. Maas M, Lambregts DM, Nelemans PJ, et al. Assessment of clinical complete response after chemoradiation for rectal cancer with digital rectal examination, endoscopy, and MRI: selection for organ-saving treatment. *Ann Surg Oncol*. 2015;22:3873–3880.

10. Van der Sande ME, Beets GL, Hupkens BJP, et al. Response assessment after (chemo)radiotherapy for rectal cancer: why are we missing complete responses with MRI and endoscopy? *Eur J Surg Oncol*. 2019;45:1011–1017.
11. Habr-Gama A, Perez RO, Wynn G, et al. Complete clinical response after neoadjuvant chemoradiation therapy for distal rectal cancer: characterization of clinical and endoscopic findings for standardization. *Dis Colon Rectum*. 2010;53:1692–1698.
12. Hupkens BJP, Maas M, Martens MH, et al. Organ preservation in rectal cancer after chemoradiation: should we extend the observation period in patients with a clinical near-complete response? *Ann Surg Oncol*. 2018;25:197–203.
13. Quirke P, Dixon MF. The prediction of local recurrence in rectal adenocarcinoma by histopathological examination. *Int J Colorectal Dis*. 1988;3:127–131.
14. Chadi SA, Malcomson L, Ensor J, et al. Factors affecting local regrowth after watch and wait for patients with a clinical complete response following chemoradiotherapy in rectal cancer (InterCoRe consortium): an individual participant data meta-analysis. *Lancet Gastroenterol Hepatol*. 2018;3:825–836.
15. Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. *Radiology*. 1983;148:839–843.
16. Hiotis SP, Weber SM, Cohen AM, et al. Assessing the predictive value of clinical complete response to neoadjuvant therapy for rectal cancer: an analysis of 488 patients. *J Am Coll Surg*. 2002;194:131–135. discussion 135–136.
17. Kawai K, Ishihara S, Nozawa H, et al. Prediction of pathological complete response using endoscopic findings and outcomes of patients who underwent watchful waiting after chemoradiotherapy for rectal cancer. *Dis Colon Rectum*. 2017;60:368–375.
18. Nahas SC, Rizkallah Nahas CS, Sparapan Marques CF, et al. Pathologic complete response in rectal cancer: can we detect it? Lessons learned from a proposed randomized trial of watch-and-wait treatment of rectal cancer. *Dis Colon Rectum*. 2016;59:255–263.
19. Smith FM, Wiland H, Mace A, et al. Clinical criteria underestimate complete pathological response in rectal cancer treated with neoadjuvant chemoradiotherapy. *Dis Colon Rectum*. 2014;57:311–315.
20. Liu S, Zhong GX, Zhou WX, et al. Can endorectal ultrasound, MRI, and mucosa integrity accurately predict the complete response for mid-low rectal cancer after preoperative chemoradiation? A prospective observational study from a single medical center. *Dis Colon Rectum*. 2018;61:903–910.
21. Smith FM, Chang KH, Sheahan K, et al. The surgical significance of residual mucosal abnormalities in rectal cancer following neoadjuvant chemoradiotherapy. *Br J Surg*. 2012;99:993–1001.
22. Lambregts DMJ, Boellaard TN, Beets-Tan RGH. Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: a pictorial review. *Insights Imaging*. 2019;10:15.
23. Beets-Tan RGH, Lambregts DMJ, Maas M, et al. Magnetic resonance imaging for clinical management of rectal cancer: updated recommendations from the 2016 European Society of Gastrointestinal and Abdominal Radiology (ESGAR) consensus meeting. *Eur Radiol*. 2018;28:1465–1475.
24. Perez RO, Habr-Gama A, Pereira GV, et al. Role of biopsies in patients with residual rectal cancer following neoadjuvant chemoradiation after downsizing: can they rule out persisting cancer? *Colorectal Dis*. 2012;14:714–720.
25. Tang JH, An X, Lin X, et al. The value of forceps biopsy and core needle biopsy in prediction of pathologic complete remission in locally advanced rectal cancer treated with neoadjuvant chemoradiotherapy. *Oncotarget*. 2015;6:33919–33925.
26. East JE, Vleugels JL, Roelandt P, et al. Advanced endoscopic imaging: European Society of Gastrointestinal Endoscopy (ESGE) technology review. *Endoscopy*. 2016;48:1029–1045.