Prospective assessment of imaging after preoperative chemoradiotherapy for rectal cancer

Fabio Pomerri, MD, Salvatore Pucciarelli, MD, Isacco Maretto, MD, Michela Zandonà, MD, Paola Del Bianco, MSc, Luca Amadio, MD, Massimo Rugge, MD, Donato Nitti, MD, and Pier C. Muzzio, MD

Background. The aim of the study was to assess the accuracy of imaging techniques in predicting pathologic tumor (ypT), node (ypN) stages and the circumferential resection margin (ypCRM) status of rectal cancers after preoperative chemoradiotherapy (CRT).

Methods. Using pelvic computed tomography (CT), magnetic resonance imaging (MRI), and endorectal ultrasound (ERUS), 90 consecutive patients with locally advanced mid-to-low rectal cancer were prospectively assessed. Postirradiation T and N stages and infiltration of the CRM, as assessed by CT, MRI and ERUS, were compared with histopathologic findings.

Results. The accuracy of ypT staging was low, whatever the imaging technique used (37% by CT, 34% by MRI, and 27% by ERUS), the most frequent inaccuracy being overstaging. Imaging showed a good specificity and good negative predictive values (NPV) when mural staging was grouped into ypT ≤ 3 and ypT4 categories; in particular, ERUS achieved a 92% specificity and 95% NPV. CRM involvement was correctly predicted in 71% of patients by CT (74% specificity; 93% NPV) and in 85% by MRI (88% specificity; 95% NPV). The accuracy for nodal staging was 62%, 68%, and 65% by CT, MRI and ERUS, respectively; the corresponding NPV were 88%, 78%, and 76%.

Conclusion. Current imaging techniques are inaccurate in restaging rectal cancer after CRT but are useful in predicting T ≤ 3 tumors, cases with negative nodes and tumor-free CRM. These findings may be of clinical relevance for planning less invasive surgery. (Surgery 2011;149:56-64.)
rectal cancer. Up to 18% of patients are over- 
staged, however, and given unnecessary CRT, and as many as 23% of patients considered node negative at initial staging reveal nodal metastases at histopathology. Moreover, a more conservative approach (wait and see or local excision alone) has been recommended for patients with a major clinical response after CRT. Although the accuracy of pretreatment staging is important in selecting patients likely to benefit from CRT, the accuracy of patient restaging after CRT may be relevant to the choice of the most appropriate operative procedure. Many studies on this topic grouped patients treated with surgery alone together with those given CRT, or compared the diagnostic images obtained before CRT with histopathologic findings, thus potentially interfering with the reliability of any assessment of the efficacy of imaging in rectal cancer staging and restaging. Studies on the performance of imaging modalities after CRT also report poor results in terms of accuracy owing to the tendency to misinterpret the inflammation and fibrosis induced by the CRT, in both the pelvic tissues and the tumor. Despite this limited accuracy, some information may be clinically relevant and help in the planning of the subsequent operations: For example, a negative predictive value (NPV) regarding lymph node status may orient the choice toward a less invasive surgery (local excision instead of conventional radical surgery) after CRT, avoiding any unnecessary total mesorectal excisions.

The aim of the present study was to prospectively investigate the performance of CT, MRI, and ERUS in restaging patients with rectal cancer after CRT and before surgery, focusing on the potential impact of restaging on the subsequent operative approach.

PATIENTS AND METHODS

Patient selection. The present prospective study was conducted after obtaining approval from the local institutional review committee and in compliance with the Helsinki Declaration. Fully informed consent was obtained in writing from all patients prior to their enrolment. The following criteria were necessary for enrollment in the study: (1) Biopsy–proven adenocarcinoma up to 11 cm from the anal verge identified by rigid proctoscopy; (2) age ≤75 years; (3) an Eastern Cooperative Oncology Group performance status of 0–2; and (4) transmural and/or node-positive rectal cancer. The workup included clinical history and physical examination, proctoscopy and colonoscopy, carcinoembryonic antigen assay, thoracic, abdominal and pelvic CT, pelvic MRI, and ERUS. Patients given no CRT were excluded.

Treatments. Where discrepancies emerged between the outcomes of baseline CT, MRI, and ERUS staging, the worst stage was considered for the purpose of planning the preoperative treatment. External beam radiotherapy was delivered in fractions of 1.8 Gy/d (total dose, 50.4 Gy). Patients also received 5-fluorouracil–based chemotherapy administered in a bolus (350 mg/m² per day) or a continuous venous infusion (225–300 mg/m² per day). Restaging included pelvic CT, MRI, ERUS, and proctoscopy. Surgery, involving total mesorectal excision as described elsewhere, was planned 6–8 weeks after completing CRT.

Histopathology. Using the Quirke protocol, the surgical specimen was evaluated by the same team of pathologists, and findings were reported according to the American Joint Committee on Cancer post-CRT tumor–node–metastasis (ypTNM) classification. A pathologic complete response was defined as the absence of viable tumor cells in the surgical specimen, and pure acellular mucin was considered as “no residual tumor.” The distance between the CRM and the nearest edge of the tumor was measured and the CRM was considered positive if the distance was ≤1 mm. When lymph node metastases, tumors in lymphatics and veins, or isolated tumor deposits were closer to the CRM than the main tumor, these lesions were used to measure the distance from the CRM.

Imaging techniques. To reduce any bias relating to the time elapsed between completing CRT and surgery, restaging was scheduled between the 4th and 5th weeks after completing the CRT. In the morning of the day when CT was scheduled, all patients were given a rectal-cleansing enema and then underwent CT with a helical CT scanner (Somatom Emotion; Siemens Medical Systems, Erlangen, Germany). All patients were administered an intravenous contrast medium (2 mL/kg; flow rate 3 mL/sec; Omnipaque 350; Nycomed Imaging A.S., Oslo, Norway), and 3-mm cuts with a pitch of 1.5 were taken through the pelvis, from the anal verge to the iliac crests. Axial images were reconstructed in the coronal and sagittal planes for interactive multiplanar image viewing on a workstation. On another day, in the morning of the day when MRI was scheduled, all patients received a rectal-cleansing enema and then underwent MRI, performed with a 1.0-T unit (Magnetom Harmony; Siemens Medical Systems), using a phased array surface coil. Sagittal, axial, and coronal T2-weighted, turbo spin-echo images were obtained first. Then axial
T1–weighted, spin-echo, gadolinium-enhanced images (0.2 mL/kg; Magnevist; Schering AG, Berlin, Germany) were recorded. The transverse images were acquired at right angles through the tumor in 4-mm thick sections with a 0.4-mm gap, a 168 × 384 matrix, and a 25-cm field of view.

Real-time ERUS was performed at a frequency of 5–10 MHz using a 5- to 30-mm focal distance with a 360° radial rotating transducer and a balloon filled with degassed water at the probe tip (Leopard Ultrasound Scanner; B&K Medical, Gentofte, Denmark). Patients were given a rectal-cleansing enema in the morning of the day when ERUS was performed. CT, MRI, and ERUS were all performed on different days but within the same week.

Image analysis. The ERUS findings were analyzed by the same qualified colorectal surgeon familiar with the technique. Given the operator-dependent nature of ERUS, intra- and interobserver variability were quantified in the staging of rectal tumors using only pelvic CT and MRI. Images were analyzed separately and then jointly by 2 teams, each consisting of 3 radiologists with different experience in interpreting pelvic CT and MRI studies. Each radiologist assessed the images again ≥6 weeks later, when the cases were presented to the teams in a different order. Each radiologist knew that patients had been referred for rectal cancer restaging, but was not informed of the other radiologists’ findings and was unaware of the final operative and histopathologic results. All images were analyzed and reviewed at a workstation. CT, MRI, and ERUS images were assessed for depth of tumor infiltration and metastatic lymph nodes according to the TNM classification. For T staging, the findings were divided into 3 groups. Tumors that seemed to be confined to the rectal wall were classified as T1–T2. Full-thickness involvement of the rectal wall with nodular or grossly spiculated tumor infiltration of the perirectal fat was classed as T3. Cases with invasion of surrounding organs or structures other than the perirectal fat were recorded as T4. Lymph nodes >5 mm in their transverse diameter were considered positive for metastases. CT and MRI images were also assessed to predict CRM involvement. If the nearest distance between the mesorectal fascia and extramural tumor, suspected lymph node metastases, or extramural tumor deposits was ≤1 mm, then the CRM was considered positive. The most relevant analyses were those relating to findings with a potential influence on subsequent surgery, for example, postchemoradiation clinical tumor (ycT) stages were grouped into ycT0 and ycT1 categories, and postchemoradiation clinical node (ycN) stages into ycN0 (negative) and ycN+ (positive) categories because a less aggressive operation may be planned for patients with pathologic complete response (ycT0N0 stage). Again, ycT stages were grouped into ycT ≥1 and ycT4 categories because a multivisceral resection may be required for stage ycT4 as opposed to the standard total mesorectal excision (TME) used for stages yc ≤T3. Nodal status was classified according to the conventional 5-mm lymph node cutoff. Because a cutoff of 10 mm (in transverse diameter) has very recently been proposed as evidence of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Age, median (range, years)</th>
<th>Gender, n (%)</th>
<th>Tumor location, n (%)</th>
<th>Tumor distance from anal verge, median (range, cm)</th>
<th>Operative procedure, n (%)</th>
<th>ypT stage, n (%)</th>
<th>ypN stage, n (%)</th>
<th>ypM stage, n (%)</th>
<th>ypCRM, n (%)</th>
<th>ypTNM stage, n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>61 (20–81)</td>
<td>Male 55 (61)</td>
<td>Lower rectum 63 (70)</td>
<td>Abdominoperineal resection 10 (11)</td>
<td>0</td>
<td>14 (16)</td>
<td>Negative 66 (73)</td>
<td>0 79 (88)</td>
<td>78 (91)</td>
<td>0 14 (16)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Female 35 (39)</td>
<td>Middle rectum 27 (30)</td>
<td>Low anterior resection 80 (89)</td>
<td>1</td>
<td>2 (2)</td>
<td>Positive 24 (27)</td>
<td>11 (12)</td>
<td>8 (9)</td>
<td>1 29 (32)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2</td>
<td>10 (11)</td>
<td></td>
<td></td>
<td>8 (9)</td>
<td>21 (23)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>3</td>
<td>23 (25)</td>
<td></td>
<td></td>
<td>Not valuable 4(4)</td>
<td>15 (17)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>4</td>
<td>34 (38)</td>
<td></td>
<td></td>
<td></td>
<td>11 (12)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Low rectum, <7 cm from anal verge; mid rectum: 7–11 cm from anal verge.
\*ypTNM, Postchemoradiation pathologic tumor, node, metastasis staging system; ypCRM, postchemoradiation pathologic circumferential resection margin.
nodal involvement in patients who have had CRT, an unplanned additional analysis was performed using the 10-mm cutoff and data were reported along with those obtained using the original 5-mm cutoff.

**Statistics.** Observer agreement for CT and MRI was calculated using Cohen’s kappa test. Values between 0 and 0.20 corresponded with slight agreement, between 0.21 and 0.40 with fair agreement, between 0.41 and 0.60 with moderate agreement, between 0.61 and 0.80 with substantial agreement, and between 0.81 and 1.0 with near-perfect agreement. CT, MRI, and ERUS postirradiation staging was compared with the corresponding pathologic T and N, and CRM status. The following were calculated: Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and accuracy, with 95% confidence intervals. Statistical significance was set at \( P \leq .05 \). The statistical analyses were carried out with SAS statistical software (rel. 8.02; SAS Institute, Cary, NC).

**RESULTS**

**Patients and treatments.** Ninety patients met the inclusion criteria and formed the study group, which included 46 patients who have been described in a previous article. Patient, tumor, and treatment characteristics are summarized in the Table. Of the 90 patients restaged after CRT, 79, 68, and 83 underwent pelvic CT, MRI, and ERUS, respectively. The remainder lacked post-CRT imaging findings for the following reasons: CT was refused \( (n = 9) \) or contrast-enhanced CT was contraindicated \( (n = 2) \); MRI was refused owing to claustrophobia \( (n = 7) \) or not performed for technical reasons \( (n = 15) \); ERUS was refused \( (n = 2) \); or the passage of the US transducer was not tolerated \( (n = 5) \). Overall, all 3 imaging modalities were used in the restaging of 53 patients, 2 modalities in 34 cases and only 1 modality in 3.

**Observer agreement.** CT and MRI afforded from substantial to near-perfect intra- and interobserver agreements in detecting tumor depth, grouping mural stages into \( ycT \leq 3 \) and \( ycT 4 \) categories (mean kappa values, 0.66–0.89), and substantial intra- and interobserver agreement in evaluating tumor involvement of the CRM (mean kappa values, 0.61–0.81) and nodal status (mean kappa values, 0.66–0.81).

**T stage.** The distribution of the ypT stage is summarized in the Table. Two ypT0 patients had lymph node metastases. When ycT and ypT stages were compared, the accuracy for CT, MRI, and ERUS was 37%, 34%, and 27%, respectively. The most frequent inaccuracy related to overstaging, which occurred in 48%, 43%, and 42% of cases using CT, MRI, and ERUS, respectively. The corresponding figures for understaging were 15%, 24%, and 31%. Of note, most ypT4 tumors understaged by CT \( (3 \text{ of } 5) \), MRI \( (4 \text{ of } 6) \), and ERUS \( (4 \text{ of } 5) \) had minimal microscopic foci of tumor
deposits in the vaginal posterior wall, peritoneal reflection, pelvic side wall, or cervix.

Figure 1 illustrates the difficulty encountered in interpreting extramural invasion after neoadjuvant therapy. The sensitivity, specificity, PPV, NPV, and overall accuracy for each imaging modality, grouping mural stages into ycT0 versus ycT $\leq$1, or ycT $\geq$3 versus ycT4, are summarized in Fig 2, A. The noteworthy finding in this figure is the very low specificity, meaning a poor prediction of any pathologic complete response. It is also worth noting the high specificity and NPV on using all 3 imaging techniques to distinguish between ycT $\geq$3 and ycT4 cases, although the small number of patients in stage ypT4 (7 of 90) may have affected these results. These results have been confirmed in the group of 53 patients who underwent all the 3 modalities (Fig 2, B).

N stage. The distribution of the ypN stage is summarized in the Table. Patients with pathologic node involvement amounted to 24 out of 90 (27%). When compared with the final histopathologic evaluation, postirradiation CT, MRI, and ERUS assessment accurately predicted N stage (using a cutoff of 5 mm) in 62%, 68%, and 65% of cases, respectively. CT, MRI, and ERUS prompted the overstaging of 32%, 18%, and 17% of patients, respectively. The corresponding figures for understaging were 6%, 15%, and 18%, respectively.

Figure 3, A shows how the imaging modalities performed using a cutoff of 5 or 10 mm for nodal involvement. When the 10-mm cutoff was used, the accuracy rate was 82% for CT and 75% for MRI, the specificity was 94% for both CT and MRI, and the NPV was 85% for CT and 76% for MRI. These results have been confirmed in the group of 53 patients who underwent all the 3 modalities (Fig 3, B).
CRM involvement. CRM was evaluated in 86 patients: 78 were CRM negative and 8 were CRM positive. In 4 cases, CRM involvement was not reported on histopathologic examination because the mesorectal excision was incomplete. CRM involvement was correctly identified in 71% and 85% of patients who underwent post-CRT CT and MRI, respectively. The most frequent inaccuracy was overstaging, in 24% of cases by CT and 11% by MRI.

Figure 4 illustrates the difficulty encountered in interpreting CRM status after neoadjuvant therapy. Figure 5, A compares the overall accuracy, sensitivity, specificity, PPV, and NPV of CT and MRI in identifying CRM involvement and shows that imaging effectively identifies a tumor-free CRM, although here again the small number of CRM-positive cases (15/65) may have affected our results. These results have been confirmed in the group of 53 patients who underwent all the 3 modalities (Fig 5, B).

DISCUSSION

This prospective study was conducted to evaluate the performance of the most frequently used imaging techniques in staging rectal cancer patients after CRT. Although the overall accuracy of all imaging techniques was poor, the major finding of this study was that CT and MRI were able to exclude advanced T stages (ypT4) and CRM involvement, and they showed a high specificity and NPV for N staging. In patients with locally advanced rectal cancer, CRT can induce tumor downsizing and downstaging. A better pathologic response to such treatment is usually associated with a better outcome,31,32 and a wait-and-see approach10 or local resection (ie, transanal local excision)11-13,32 have been advocated in patients with a major clinical response. It is consequently important to improve the accuracy of restaging after CRT, particularly to rule out any residual cancer in the rectal wall or mesorectum. Unfortunately, because it is difficult to distinguish residual cancer from postirradiation peritumoral vasculopathy and desmoplastic reactions, the accuracy of restaging has proven unsatisfactory. As reported by others,14,19-22,33 our findings confirm that current imaging techniques are unreliable in predicting ypT stage. The high specificity and NPV in predicting ypT \( \leq 3 \) and ypT4 cases that we identified may be useful for clinical purposes, but our small number of patients with ypT4 may be a bias, and therefore no firm conclusions can be drawn on this issue. On the whole, our findings do not support the “wait-and-see policy” recommended by some authors for patients with a clinically complete response after CRT.10

Accuracy in predicting metastatic lymph nodes in the mesorectum is even more important than in predicting T stage because the main concern when performing a local excision after CRT is the risk of leaving nodal disease in the mesorectum. The best criterion to use in predicting metastatic lymph nodes remains a matter of debate. In addition to the dimensional criterion, the morphologic characteristics of the lymph node (ie, contours and
textural features) have also been considered as predictors of malignancy. Based on previous studies on patients not given CRT, we used the 5-mm cutoff as a dimensional criterion for predicting lymph node involvement. We are aware of the weakness of this approach, but it is still the simplest and most widely used in clinical practice. It is likely that CRT affects all previously used criteria (size, shape, and texture), and the same authors recently found that the best node size cutoff for predicting nodal involvement after CRT was 10 mm, not 5 mm. In our study, using the 10-mm cutoff gave rise to an accuracy of 82% for CT and 75% for MRI, a specificity of 94% for both CT and MRI, and an NPV of 85% for CT and 76% for MRI. Specificity and NPV are of the utmost relevance when a local excision procedure is planned after CRT, because a false-negative diagnosis of nodal metastases would lead to the mesorectum not being removed, which raises much more concern than a false-positive one (which would prompt total mesorectal excision in patients with no nodal metastases). Although better results might be achievable in the future by using more advanced technologies, the simultaneous use of multiple diagnostic criteria and markers for lymph node involvement, the dimensional criterion adopted in the current study may still have a role in clinical practice because it is simple and readily available worldwide.

CRM involvement has been considered a strong predictor of recurrence in cases of rectal cancer, both in patients given no CRT and in those given short-course radiotherapy. More recently, CRM involvement has been used to plan postoperative CRT. Little is known about the accuracy of imaging in defining CRM involvement after CRT. In our study, MRI revealed a high accuracy (85%) in predicting CRM involvement, in line with 2 recent studies reporting an accuracy of 77% and 82%, albeit at the cost of a false positive rate of around 50%. We found a higher false-positive rate (70%) using MRI. The imbalance between the proportions of positive- and negative-CRM cases may have contributed to the discrepancy found between our findings and those reported by other authors. The high proportion of true negative CRM cases in our cohort of patients (52/65) probably accounts for the high specificity (88%), and our relatively high accuracy (85%) may have depended on the small proportion of positive cases (3/65). On analyzing histopathologic data for non-irradiated patients, some authors have found a significantly lower local recurrence rate in cases with a CRM >2 mm than in those with a margin of 1–2 mm. As in recent reports on irradiated patients, we chose the classic distance of the tumor from the mesorectal fascia of 1–2 mm.

Although our findings on the accuracy of CT and MRI were disappointing, we found that these techniques have a high NPV for CRM involvement. In particular, our NPV results on MRI were similar to those reported in previous studies evaluating tumor-free margins in irradiated patients, irrespective of whether a cutoff of 2 or 1 mm was used. Observer agreement measures provide information on the reliability of a diagnosis based on imaging techniques, but they do not reflect the diagnostic accuracy of pelvic CT and MRI. In the present study, despite the radiologists’ different levels of training and confidence in interpreting post-CRT images, a substantial observer agreement

![Figure 5](image-url)
emerged for the depth of tumor invasion and for CRM and regional lymph node involvement. Although intra- and interobserver agreement was acceptable, the comparison drawn between CT/MRI and the histopathologic findings showed that these widely used imaging techniques are inaccurate in evaluating T stage and lymph node status after CRT. The reason for the poor correlation between histopathologic and imaging findings consequently seems to depend more on the objective difficulty involved in interpreting the images rather than on observer performance.

Tumor response to therapy can be assessed noninvasively using functional imaging by positron emission tomography (PET) with $^{18}$F-2-deoxy-D-glucose (FDG) or a scanner combining PET with CT technology (FDG-PET/CT). Some authors have suggested that FDG-PET may be more accurate than other imaging modalities in detecting a tumor response, but the technique has its weaknesses. Interpretational errors might stem from no increase in tracer uptake in responding tumors with residual areas of disease or in cases of a reduction in glucose metabolism after radiotherapy. Moreover, poorly visible tumors in perirectal lymph nodes might be responsible for an inaccurate assessment of lymph node status. FDG-PET/CT is also of little use in predicting tumor-free mesorectal fascia.

The present study has some limitations. The imaging techniques discussed are no longer considered the state of the art. Although ordinary 2-dimensional imaging data may be used for multiplanar image reconstruction and viewing, recent technical advances (in terms of an increased speed and spatial resolution) have enabled an improvement in CT and MRI image postprocessing quality. Although results tend to be technology dependent, authors using modern methods (eg, 1.5-T MRI and multidetector row CT) have reported a low to moderate accuracy of imaging for rectal cancer restaging after preoperative neoadjuvant therapy. The clinical relevance of the techniques used in the present study lies mainly in that CT and MRI achieved a good specificity and NPV in predicting rectal cancer stage (classifying the tumors as $T_≤3$ and $T_4$), lymph node status (using a cutoff of 10 mm), and circumferential margin invasion (using a distance of the tumor from the mesorectal fascia of $≤1$ mm). Second, our series included a small number of $T_4$ lesions (8%) and cases with CRM involvement (9%). Therefore, how well our observations regarding the performance of imaging apply to patients with $T_4$ disease or positive CRM is hard to say.

In conclusion, most CT, MRI, and ERUS findings have no impact on the therapeutic outcome of chemoradiation-treated patients with locally advanced rectal cancer, owing to the poor consistency between post-CRT imaging and histopathologic staging. CT, MRI, and ERUS have a marked ability to predict $T_≤3$ status, tumor-free CRM, and lymph node negativity after CRT, but their poor performance in the local restaging of chemoradiation-treated cases of mid-to-low rectal cancer suggests that currently used restaging criteria should be reconsidered in an attempt to maximize accuracy and minimize the high false-positive rate encountered in the assessment of $T$ stage and lymph node status.

REFERENCES


