Magnetic resonance imaging in locally advanced rectal cancer: quantitative evaluation of the complete response to neoadjuvant therapy

Nicola Tarallo B,D,E, Maria Gloria Angeretti A,B,D, Elena Bracchi A,B,D,E,F, Genti Xhepa B,D,E, Valeria Molinelli A,B,F, Chiara Tagliaferro B,D,E,F, Paolo Antognoni B,D,E, Raffaele Novario C,D, Fausto Sessa B,D, Carlo Fugazzola A,D,E,F

Università degli Studi dell’Insubria – Circolo Hospital and Macchi Foundation, Italy

Abstract

Purpose: To assess the diagnostic performance of diffusion-weighted imaging (DWI) for the discrimination of complete responder (CR) from the non-complete responder (n-CR) in patients with locally advanced rectal cancer (LARC) undergoing chemotherapy and radiation (CRT).

Material and methods: Between December 2009 and January 2014, 32 patients (33 lesions: one patient had two synchronous lesions) were enrolled in this retrospective study. All patients underwent a pre- and post-CRT conventional MRI study completed with DWI. For both data sets (T2-weighted and DWI), the pre- and post-CRT tumour volume (VT2; VDWI) and the tumour volume reduction ratio (∆V%) were determined as well as pre- and post-CRT apparent diffusion coefficient (ADC) and ADC change (∆ADC%). Histopathological findings were the standard of reference. Receiver operating characteristic (ROC) curves were generated to compare performance of T2-weighted and DWI volumetry, as well as ADC.

Results: The area under the ROC curve (AUC) revealed a good accuracy of pre- and post-CRT values of VT2 (0.86; 0.91) and VDWI (0.82; 1.00) as well as those of ∆VT2% (0.84) and ∆VDWI% (1.00) for the CR assessment, with no statistical difference. The AUC of pre- and post-CRT ADC (0.53; 0.54) and that of ∆ADC% (0.58) were significantly lower.

Conclusions: Both post-CRT VDWI and ∆VDWI% (AUC = 1) are very accurate for the assessment of the CR, in spite of no significant differences in comparison to the conventional post-CRT VT2 (AUC = 0.91) and ∆VT2% (AUC = 0.84). On the contrary, both ADC and ∆ADC% values are not reliable.

Key words: magnetic resonance imaging, diffusion-weighted imaging, rectal cancer, chemoradiation therapy.

Introduction

Locally advanced rectal cancer (LARC) – defined by the parameters cT3-T4, N –/+, and M0 – has a five-year survival rate equal to 50-65%, with a local recurrence rate of 30-40% and a high incidence of metastasis [1,2]. Over the past 20 years, the “total mesorectal excision” (TME) has fostered a considerable reduction of recurrence of LARC; nevertheless, it has significant post-surgical complications (anorectal, urinary, and sexual dysfunction), and post-operative mortality is still relevant (range: 2-30% within six months of surgery) [3,4].

Currently the treatment of LARC is oriented towards the use of neoadjuvant therapies, radiotherapy combined with chemotherapy (CRT), which further reduces the recurrence rate, leading to volume reduction and tumour downstaging [4]. In patients affected by LARC, a complete histopathological response after CRT occurs in 10-30% of cases at the time of surgery and is related with better prognosis in terms of local recurrence as well as disease-free
and overall survival; it has also been demonstrated that neoadjuvant therapy alone, in an accurately selected group of patients, represents a safe treatment associated with a good survival rate at five years (overall survival 88% and disease-free 83%) [2,4-7]. A “wait and see” approach has been proposed for patients with complete clinical response after CRT. For this reason, it is necessary to identify imaging methods capable of discriminating “complete responder” patients (CR) from “non-complete responder” patients (n-CR) [2,5-7].

Magnetic resonance imaging (MRI) is the best imaging technique for the evaluation of tumour response to CRT [8,9]. However, the conventional MR sequences are not sufficiently reliable in distinguishing between residual tumour and fibrotic reaction after chemoradiation [1,8,9].

The current guidelines of the European Society of Gastrointestinal and Abdominal Radiology (ESGAR) suggest the use of diffusion-weighted MR imaging (DWI) in post-CRT restaging [8] because there is growing evidence that the qualitative analysis of DWI sequences improves the diagnostic performance in the evaluation of the response (stage-yT) to neoadjuvant CRT [4,8-10]. Diffusion-weighted MRI is a functional imaging technique that analyses differences in the extracellular movement of water protons to discriminate between tissues of varying cellularity, with a potentially beneficial role for the detection and characterisation of malignant tumours; it can provide a quantitative assessment by measuring the apparent diffusion coefficient (ADC) [5,11-21]. In order to identify CR patients, a few studies have performed quantitative assessments by calculating pre- and post-CRT mean ADC values of the tumour [12,15-17,22], as well as the percentage of ADC change (ΔADC%) [12,15,17], with contradictory results.

Other studies, by calculating the tumour volume on conventional MR sequences (V_T2) pre- and post-CRT and the volume reduction rate (ΔV_T2%), [23-26], have also obtained conflicting results. Therefore, in order to distinguish CR from n-CR, other authors have proposed the evaluation of the functional volumetry (V_DWI) performed on pre- and post-CRT DW images on high b-value and the volume reduction rate (ΔV_DWI%) [5,14,27].

The aim of our study was to assess the diagnostic performance of DWI by calculating the V_DWI (compared to V_T2), the ADC values, and the relative Δ% for the discrimination of CR from the non-CR after CRT.

Material and methods

Fifty patients with MRI diagnosis of LARC between December 2009 and January 2014 were considered for inclusion in our retrospective study based on the following criteria:

- endoscopic diagnosis and histopathologically (biopsy) proven rectal carcinoma;
- conventional MR pre-CRT completed with DWI which confirmed the tumour and allowed accurate locoregional staging (all tumours were T3 – locally advanced – regardless of the distance from the mesorectal fascia);
- combined neoadjuvant therapy: the treatment protocol included external beam radiotherapy for a total of 45 to 50.4 Gy (1.8 Gy/fraction) and chemotherapy with 5-fluorouracil (continuous infusion of 225/mg/m²/day for seven days for the duration of radiation therapy) or Capecitabine per os (825 mg/m² two times per day, from Monday to Friday, for the duration of the radiation treatment);
- conventional MR completed with DWI after neoadjuvant treatment;
- histopathological examination of the surgical specimen or, alternatively, biopsy performed during follow-up endoscopy in patients with strong evidence of complete response to therapy based on clinical and instrumental investigations, in which an attitude of surveillance was considered preferable to surgical approach.

Of the 50 patients initially enrolled, 18 were excluded for the following reasons: two patients for metastatic disease and comorbidities; one patient for poor quality of DWI due to artefacts caused by metallic hip implants; four patients lost at follow-up (FU) after performing post-CRT MR; and 11 patients underwent surgery after staging MRI. The final population eligible for our study encompassed 32 patients (33 lesions in 32 patients: one patient had two synchronous lesions, one in the rectum and one in the anal canal): 18 males and 14 females – mean age 65.9-years (range: 35-85 years).

All MR images were retrospectively evaluated in consensus by two radiologists (GA; EB); the observers were blinded to the clinical patient data and pathology reports.

Standard of reference

Twenty-nine of 32 patients underwent TME; 3/32 patients did not undergo surgery, due to strong clinical evidence of a complete response (repeated negative colonoscopy and biopsies after CRT).

Tumour response after CRT was determined in all the 33 lesions according to the pathological classification suggested by Dworak’s tumour response grading system [14,28] (TRG): grade 4, no tumour cells, only a fibrotic mass (CR); grade 3, very few tumour cells in fibrotic tissue with or without mucous substance (near CR); grade 2, dominantly fibrotic changes with few minor cells or groups (moderate response); grade 1, dominant tumour mass with obvious fibrosis and/or vasculopathy (minimal response); grade 0, no regression.

MRI technique

All patients provided written, informed consent and were investigated by MRI with a magnetic field of 1.5 Tesla.
by the section thickness and then summing all the partial volumes [5,27,29] (Figure 1A, B, D, E).

For both data sets (T2 weighted and DWI), the pre- and post-CRT tumour volumes (V\text{pre} and V\text{post}) were determined; moreover, the tumour volume reduction ratios for both T2-weighted and DW images (ΔV\text{pre} and ΔV\text{post}) were calculated as follows: (V\text{pre} − V\text{post})/V\text{pre} × 100 [5,14,27].

**Measurement of the ADC**

ADC maps were automatically generated by using a monoexponential decay model including the two b values (0 and 800 s/mm\(^2\)), on which freehand ROIs were drawn at the edges of the tumour for each axial section containing the lesion [5,29] (Figure 1C).

The tumour margins on ADC maps were defined referring to the paraxial T2-weighted and DW images on the high b value (b = 800 s/mm\(^2\)): mean ADC value was extrapolated by ADC values obtained in the axial sections and the relative standard deviations with the goal of reducing the structural differences induced by the inherent tumour heterogeneity.

When no remaining high SI was visualised on the post-CRT DW images (Figure 1E), three ROIs were drawn at the former location of the primary tumour with reference to the post-CRT paraxial T2-weighted images [5,29] (Figure 1F).

Mean ADC values of the tumour lesions (pre- and post-CRT) as well as the percentage of ADC change (ΔADC%) were calculated. ΔADC% was determined as follows: (ADC\text{post} − ADC\text{pre})/ADC\text{pre} × 100 [15].

**Statistical analysis**

The following statistical analyses were performed by using GraphPad software (MedCalc Software, version 5.02):
- the Wilcoxon test was used to compare pre- vs. post-CRT V\text{pre} vs. V\text{post} and mean ADC values of the 33 lesions included in the study;
- the Mann-Whitney test was used to compare pre- vs. post-CRT V\text{pre} vs. V\text{post} and mean ADC values, and relative ratio (ΔV\text{pre} vs. ΔV\text{post} %; ΔADC%) for CR vs. n-CR;
- receiver operating characteristic (ROC) curves were generated and applied to the above values to assess the diagnostic performance in detecting a CR; corresponding area under the ROC curve (AUC) was considered for the determination of diagnostic accuracy and a 95% confidence interval (CI) expressed the statistical precision of the results. For these analyses, cut-off values were determined according to the point nearest to the upper left corner in the ROC curves. The interpretation of the AUC values refers to the classification proposed by Swets [30]. A difference with a P value < 0.05 was considered significant.
MR imaging in locally advanced rectal cancer

Figure 1. 61-year-old woman with distal rectal advanced adenocarcinoma: CR post-CRT. A) Pre-CRT T2-weighted axial image shows the mass; freehand ROI was drawn along the border of the lesion for calculation of the sectional area of tumour; sectional area was multiplied by section thickness to determine the tumour volume \( (V_{T2} = 6.8 \text{ cm}^3) \). B) Pre-CRT DW axial image (\( b \) value: 800 s/mm\(^2\)): a freehand ROI was drawn for the calculation of the sectional area of tumour and of the tumour volume (\( V_{DWI} = 1.97 \text{ cm}^3 \)). C) Pre-CRT axial ADC map; a freehand ROI was drawn for the calculation of the ADC; the mean ADC value was calculated from the different axial values (\( 1.23 \pm 0.26 \times 10^{-3} \text{ mm}^2/\text{s} \)). D) Post-CRT T2-weighted axial image shows slight rectal wall thickening with hypointense signal, interpreted as fibrosis; freehand ROI was drawn along the border of the thickening; sectional area was multiplied by section thickness to determine the tumour volume \( (V_{T2} = 0.94 \text{ cm}^3) \) and the tumour reduction ratio \( (\Delta V_{T2\%} = 86.2\%) \). E) Post-CRT DW axial image (\( b \) value: 800 s/mm\(^2\)): no residual hyperintense signal is observed in the corresponding lesion, therefore tumour volume \( (V_{DWI}) \) is equal to 0 cm\(^3\) and \( \Delta V_{DWI\%} \) to 100%. F) Post-CRT axial ADC map; freehand ROI is drawn along former location of the tumour demonstrated by the T2 weighted image, for the calculation of the ADC (mean value: \( 1.73 \pm 0.19 \times 10^{-3} \text{ mm}^2/\text{s} \)) and ADC change \( (\Delta ADC\% = 41\%) \). A correct prediction of the complete tumour response was made by post-CRT \( V_{T2} \) and \( V_{DWI} \), as well as by \( \Delta V_{DWI\%} \). Pathological examination of resected specimen revealed no residual tumour cells (TRG 4).

Results

Fourteen of the 32 patients included in our study underwent anterior resection of the rectum and 15/32 underwent abdominoperineal resection. Three patients were monitored with endoscopy and concomitant biopsy (negative results for the presence of residual disease) and underwent FU with MRI investigation every 3-6 months for the first year and then annually (mean FU: 5 years and three months).

The mean time between the end of neoadjuvant therapy and restaging MR imaging was 53.4 days (range: 38-82 days), and the mean time between the post-CRT MR imaging and surgery (or biopsy) was 21.2 days (range: 2-60 days).

Pathologic examination revealed 32/33 rectal adenocarcinoma and 1/33 mucinous type adenocarcinoma (32 patients, one with a double synchronous lesion). The locations of the tumours were: four proximal rectum, 17 middle rectum, eight distal rectum, and four anal canal. The histopathological examination provided the following results: 4/33 lesions were considered G0, 13/33 as G1, 7/33 as G2, 2/33 as G3, and 7/33 as G4 (4/7 with histopathological examination of the surgical specimen and 3/7 evaluated on the biopsy material); the CR group was therefore composed of 7/33 patients and the n-CR group of 26/33 patients. In relation to the (yT) staging: 7/33 lesions were yT0 (including 4ypT0); 2/33: ypT1; 9/33: ypT2; 12/33: ypT3; 3/33: ypT4.

Conventional T2-weighted volumetry vs. DWI volumetry

A statistically significant \( (p < 0.0001) \) reduction of median tumour volume of both \( V_{T2} \) and \( V_{DWI} \) Pre- vs. post-CRT was noted for all 33 lesions included in the study, respectively, from 26.4 cm\(^3\) to 11.4 cm\(^3\) and from 14.8 cm\(^3\) to 5.3 cm\(^3\) (Table 1).

The pre- and post-CRT median tumour volumes in the CR group were significantly lower compared to the n-CR group, both on T2-weighted images (\( V_{T2} \) pre-CRT: 16.1 cm\(^3\) vs. 29.97 cm\(^3\), \( p = 0.0037 \); \( V_{T2} \) post-CRT: 3.1 cm\(^3\) vs. 14.3 cm\(^3\), \( p = 0.001 \)) and on DWI (\( V_{DWI} \) pre-CRT: 6.6 cm\(^3\) vs. 17.99 cm\(^3\), \( p = 0.008 \); \( V_{DWI} \) post-CRT: 0.00 cm\(^3\) vs. 8.7 cm\(^3\), \( p = 0.0001 \)) (Table 1). The \( \Delta V\% \) was significantly higher in the CR group compared to n-CR group, both in T2-weighted images (\( \Delta V_{T2\%} \): 84.9% vs. 50.7%, \( p = 0.005 \)) and in DWI (\( \Delta V_{DWI\%} \): 100% vs. 43.7%, \( p = 0.0001 \)) (Table 1).
Table 1. Median volumes and ADC values (33 tumours)

<table>
<thead>
<tr>
<th>Measurements</th>
<th>Lesions (n = 33)</th>
<th>CR (n = 7)</th>
<th>n-CR (n = 26)</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>VT2 pre-CRT (cm³)</td>
<td>26.4 (2.2-101.6)</td>
<td>16.1 (2.2-24.9)</td>
<td>29.97 (8.5-101.6)</td>
<td>0.0037</td>
</tr>
<tr>
<td>VT2 post-CRT (cm³)</td>
<td>11.4 (0.6-93.9)</td>
<td>1.3 (0.6-11.4)</td>
<td>14.3 (1.3-93.9)</td>
<td>0.001</td>
</tr>
<tr>
<td>ΔVT2%</td>
<td>49.5 (1.5-94.9)</td>
<td>84.9 (49.5-94.9)</td>
<td>50.7 (1.5-88.6)</td>
<td>0.005</td>
</tr>
<tr>
<td>VDWI pre-CRT (cm³)</td>
<td>14.84 (1.97-101.3)</td>
<td>6.6 (1.97-17.5)</td>
<td>17.99 (3.7-101.3)</td>
<td>0.0082</td>
</tr>
<tr>
<td>VDWI post-CRT (cm³)</td>
<td>5.3 (0.6-92.9)</td>
<td>0.00 (0-0.5)</td>
<td>8.7 (1.16-92.9)</td>
<td>0.0001</td>
</tr>
<tr>
<td>ΔVDWI%</td>
<td>55.3 (2-100)</td>
<td>100 (94.2-100)</td>
<td>43.7 (2-83)</td>
<td>0.0001</td>
</tr>
<tr>
<td>ADC pre-CRT (× 10⁻³ mm²/s)</td>
<td>1.11 (0.74-1.9)</td>
<td>1.11 (0.89-1.23)</td>
<td>1.11 (0.74-1.9)</td>
<td>0.8</td>
</tr>
<tr>
<td>ADC post-CRT (× 10⁻³ mm²/s)</td>
<td>1.47 (1-2.1)</td>
<td>1.42 (1.2-1.7)</td>
<td>1.51 (1.1-2.1)</td>
<td>0.7</td>
</tr>
<tr>
<td>ΔADC%</td>
<td>24.5 (0.8-95)</td>
<td>28 (23-41)</td>
<td>22 (0.8-95)</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Median ADC value

For the 33 lesions included in the study a significant increase in median ADC value post- vs. pre-CRT (1.47 ± 0.27 × 10⁻³ mm²/s vs. 1.11 ± 0.29 × 10⁻³ mm²/s, p = 0.0001), expressed by the value ΔADC% (24.5%), was found.

Nevertheless, no statistically significant difference was found between median ADC values pre-CRT (p = 0.8), post-CRT (p = 0.7), and ΔADC% (p = 0.4) in the CR group compared to the n-CR group (Table 1).

Diagnostic performance to identify complete responder

The ROC curves were used to compare the diagnostic performances of VT2, VDWI and ΔADC% vs. the ADC values, pre- and post-CRT; nevertheless, there were no significant differences in AUC among VT2, VDWI and ΔADC% (Table 2, Figure 2A, B).

The results for AUC of ΔVT2% (0.84) and ΔVDWI% (1.00) were significantly better than AUC of ΔADC% (0.58) (Table 2, Figure 2C), without statistically significant differences between the ΔVT2% and ΔVDWI%.

In the absence of significant differences between VT2 and VDWI (pre- and post-CRT) and ΔV%, the more accurate parameters for the assessment of CR (AUC = 1) were represented by the VDWI post-CRT and ΔVDWI%.

Comparison between qualitative and volumetric evaluation of DWI

We correctly identified as CR 6/7 lesions on the basis of the absence of SI on DWI on high b-value. One false neg-
Figure 2. A) Comparison of the areas under the ROC curves (AUCs) applied to the VT2, VDWI and ADC pre-CRT revealed the absence of a significant difference between the VT2 (0.86) and VDWI (0.82), both with a significant difference compared with the ADC (0.53). B) Comparison of AUCs post-CRT equally revealed the absence of a significant difference between the VT2 (0.91) and VDWI (1.00), both with a significant difference compared with the ADC (0.54). C) Comparison of AUCs applied to the respective percentage ratios (ΔVT2%, ΔVDWI% and ΔADC%) revealed absence of a significant difference between the ΔVT2% (0.84) and ΔVDWI% (1.00), both with a significant difference compared with ΔADC% (0.58). VT2 post-CRT and the ΔVDWI% were the most accurate parameters in recognising the CR (AUC = 1).

Discussion

Identifying MR predictive biomarkers or indicators of tumour response to neoadjuvant CRT in patients with LARC is still a challenge. Conventional MR sequences are not sufficiently reliable in distinguishing between residual tumour and post-CRT tissue fibrosis [1,8,9]. However, it has been shown that the qualitative assessment of DWI sequences significantly improves the diagnostic performance of conventional MRI in the evaluation of tumour response to CRT, as regards both the downstaging (T stage) [4] and the tumour response grading (TRG) [6,8,10,31]; particularly, in distinguishing between CR and n-CR patients, DWI has shown a higher sensitivity (52-64% vs. 0-40%) and an almost comparable specificity (89-97% vs. 92-98%) vs. standard MR sequences [10].

However, DW images have limitations because complete tumour regression is not always accompanied by the absence of SI due to the fact that diffuse fibrosis associated with chronic inflammation, the presence of mucin pools, the air-rectal wall interface, or the collapsed rectal wall may be visualised as high SI, making the identification of CR difficult [5,10,14]. This also occurred in 1/7 CR patients of our series (Figure 3).

Therefore, the MR quantitative evaluations have been proposed by calculating the mean ADC values [13,15-17], the conventional volumetry on T2-weighted images (VT2) [23,24] and, more recently, the volume measured on the DW images (VDWI) [5,14,27], as well as the ratio between the values before and after CRT (Δ%) [5,12,14-17,23-25,27].

Promising results have been reported regarding the calculation of the ADC to assess the response to neoadjuvant treatment in patients with complete response (TRG 4 or ypT0) as well as near-complete response (TRG 3-4 or ypT0-T2) [12-15,17,18]. When considering only studies that compared groups of patients n-CR vs. CR in relation to the TRG (G4 vs. G0-G3), some studies claim that the pre-CRT ADC values in tumours of CR patients were significantly lower than those of n-CR [12], while the post-CRT ADC values [15,22] and those of ΔADC% [12,15,17] were significantly higher. In fact, high cellularity corresponds to a good tumour response, and therefore low ADC values pre-treatment indicate the CR patients...
Figure 3. 53-year-old man with middle rectal advanced adenocarcinoma CR post-CRT. A) Pre-CRT T2-weighted axial image shows the tumoural mass bounded by freehand ROI (V_{T2} = 9.9 cm³). B) Pre-CRT DW axial image (b value: 800 s/mm²) shows an hyperintense signal in the tumoural area bounded by the freehand ROI (V_{DWI} = 8.6 cm³). C) Post-CRT T2-weighted axial MR image shows rectal wall thickening with an area of intermediate SI interpreted as residual tumour, bounded by the freehand ROI (V_{T2} = 0.9 cm³). D) Post-CRT axial DWI shows a linear hyperintense signal, interpreted as residual tumour at qualitative evaluation, bounded by the ROI (V_{DWI} = 0.5 cm³). Although both T2-weighted and DW images wrongly demonstrate a residual tumour, a correct prediction of the complete tumour response is made by post-CRT V_{T2} and V_{DWI} as well as by ΔV_{T2} % (94%) and ΔV_{DWI} % (94.2%). Pathologic examination of resected specimen revealed no residual tumour cells (TRG 4).

[18]. However, there are aggressive tumours, characterised by predominantly coagulative necrosis, and thus with low ADC values, which may not respond favourably to neoadjuvant CRT [18]. This justifies the low ADC values pre-CRT also detected in the n-CR of our series; the biological aggressiveness of these tumours is confirmed by the tendency to infiltrate the mesorectal fascia and the lymph node involvement [21].

Some studies sustain that post-CRT ADC might not be able to highlight microscopic residual tumour in areas of fibrosis and/or post-actinic necrosis and therefore does not allow reliable discrimination of CR patients from “near-CR” patients [6,15,16]. In our series, considering the entire group of lesions (CR + n-CR), the post-CRT ADC value, while being significantly higher than the pre-CRT ADC value, was not able to distinguish CR patients from n-CR; moreover, the ΔADC% was not able to make such a discrimination. Similar conclusions on the limited usefulness of the various measurements of the ADC – including the ΔADC% – for the assessment of CR also emerged from other studies [4,5,14].

Based on our experience and the literature, we believe that the conflicting results in the evaluation of the ADC can be related to the small number of patients with complete remission included in the different studies (range: 9-35 CR from a total of 30-100 CR + n-CR) [5,14-17], and to the variability of the technique employed for the measurements [32].

Concerning the volumetric evaluation with MR T2-weighted imaging, some studies associate the ability to identify the CR with a volume reduction rate (ΔV_{T2} %) > 73.6-86.6% [12,14,24-27]; in particular Curvo-Semedo...
et al. [5] reported an AUC = 0.84 and Ha et al. [14] an AUC = 0.792. However, as regards the post-CRT V_{T2}, conflicting results have been reported: Ha et al. [14] found an AUC = 0.705; on the other hand, Kang et al. did not associate the post-CRT V_{T2} with the ability to identify the CR [24]. Curvo-Semedo et al. also considered the post-CRT V_{T2} significant, although less accurate (AUC = 0.70) than ΔV_{T2}%; moreover, they did not consider the pre-CRT V_{T2} to be significant (AUC = 0.57) [5]. Lambregts et al. confirmed the results of the above-mentioned studies [5,14], reporting a good accuracy for both ΔV_{T2} (AUC = 0.78) and post-CRT V_{T2} (AUC = 0.82) as well as for the pre-CRT V_{T2} (AUC = 0.73) [27]. Finally, other studies did not find any significant difference in order to identify the CR (TRG 4) as well as the “good responders” (TRG 3-4), arguing that the pre- and post-CRT volumes measured on T2 and the respective reduction ratio could not be used for the selection of appropriate treatment options because these two parameters were not sufficiently reliable in distinguishing between residual tumour and post-CRT tissue fibrosis [23,33].

Three studies, published between 2011 and 2015, which ascribed low usefulness to the measures of the ADC in distinguishing CR patients from n-CR (as mentioned above), argued that the tumour volume measured in the DW images was more accurate than that obtained in the conventional T2 MR sequences. In particular, Curvo-Semedo et al. [5] assign the best performance to post-CRT V_{DWI} (AUC = 0.93), which is significantly more accurate than post-CRT V_{T2} (AUC = 0.70) and comparable in accuracy to both DWI (AUC = 0.92) and T2-weighted images (AUC = 0.84); conversely, the pre-CRT V_{DWI} as well as pre-CRT V_{T2} are not accurate. These findings suggest that the evaluation of post-CRT DWI can be sufficient, so pre-CRT images do not necessarily have to be assessed. Ha et al. recognised ΔV_{DWI} % as having the best performance (AUC = 0.91) in order to identify the CR; this value is significantly more accurate than ΔV_{T2} % (AUC = 0.792), as well as post-CRT ADC (AUC = 0.705). They also reported a significant difference between CR patients and n-CR with regard to the median value of post-CRT V_{DWI} calculated by Wilcoxon test (p < 0.01) [14]. Recently, Lambregts et al. [27], by using the same cut-off values proposed by Curvo Semedo et al. [5], confirmed the previous results in a bi-institutional study and pointed out that the post-CRT V_{DWI} (AUC = 0.92) has the best diagnostic performance.

The results of our series reveal good accuracy of post-CRT V_{T2} (AUC = 0.91) and ΔV_{T2} % (AUC = 0.84) (Table 2, Figure 2); post-CRT V_{DWI} and ΔV_{DWI} % results were more accurate (AUC = 1) compared to the corresponding post-CRT V_{T2} and ΔV_{T2} %; however, the differences were not statistically significant. Moreover, our values of optimal cut-off are different from those reported in the other published studies [5,14,27]. Furthermore, it has to be emphasised that even our pre-CRT V_{DWI} (AUC = 0.82) as well as pre-CRT V_{T2} (AUC = 0.86) were accurate in identifying the CR; these latter findings, in disagreement with the literature [5,14,27], show a limitation of our study, represented by the small size of the pre-CRT lesions with complete histopathological response.

Finally, all the tumour volumes measured on the DW images, as in other studies [5,14,27], were smaller than those on the T2-weighted images; on DW imaging ROIs were drawn delimiting the areas of high SI, while in T2-weighted images areas of fibrosis were also included within the ROIs because the risk of residual tumour in the fibrotic areas is known to be about 50% [5,6,14,27]. Therefore, our experience confirms that volumetry on DW images is more accurate than that on T2-weighted images: in particular, post-CRT V_{DWI} ≤ 0.5 cm^3 and ΔV_{DWI} % ≥ 83% (our values of optimal cut-off) could indicate a pathological complete response. However, it still remains difficult to differentiate between patients with a CR (TRG 4) and patients with small microscopic clusters of residual neoplasm (TRG 3) [5]; further studies are required to address this issue.

Nevertheless, at present, although the tumour volumes determined on the basis of the presence (or absence) of high-signal intensity areas on DW-MRI better represent the existence of residual viable tumour, we can hypothesise – in agreement with Curvo-Semedo [34] – that a visual evaluation of a high-signal intensity area suggestive of residual tumour is sufficient, and volumetric measurements are not even required. Such a visual approach would also be more practical and far less time consuming [34] because of the manual measurement of tumour volume; the use of a (semi-)automated segmentation approach is a potentially time-saving alternative, as demonstrated by a recent study [35]. Hence, the combined T2 morphology and qualitative/volumetric DWI evaluation can form the cornerstone of clinically applicable daily rectal MRI interpretation [36]. Moreover, the combination of MRI with clinical assessment (digital rectal examination and endoscopy) is recommended as the optional strategy for safe and accurate selection of CRs after CRT [37].

There were some limitations to our study: the small number of selected patients (especially the small sample of CR); the small size of the pre-CRT lesions with pathological complete response; histopathological evaluation of tumour regression to therapy was performed on biopsy in 3/7 CR; the lack of direct correlation between volumetric data obtained by MR images and the volumetric data provided by the surgical specimens; the possible errors in the positioning and size of the ROIs drawn on the tumour margins, not easily identifiable especially in the examinations performed after CRT, due to the low spatial resolution and low signal/noise ratio in the DW images and ADC maps; the inter-observer reproducibility of the method was not evaluated because of the long time required for measurements of volumes and ADC values; and finally, post-CRT N parameter, so far considered in a single study [4], was not assessed. However, the prevalence of a positive lymph node status in the case of CR
of the primary tumour after CRT is very low (8%) [10]; moreover, standard MRI is quite accurate in lymph node staging after CRT, so the addition of functional imaging, such as DWI, may not even be necessary [38].

In conclusion, DW images improve the results of standard follow-up MR protocols in order to identify CR patients after neoadjuvant CRT in patients affected by LARC. The functional volumetry is better than the conventional volume, although no statistically significant differences were detected in this study. In particular, both post-CRT $V_{DWI}$ and $\Delta V_{DWI}$ results are very accurate; however, standardised cut-off values are not available. Conversely, the pre- and post-CRT ADC values and $\Delta ADC$ are not sufficiently reliable to distinguish the CR patients from the total group of n-CR patients.

**Conflict of interest**

The authors report no conflict of interest.

**References**

23. Kim YH, Kim DY, Kim TH, et al. Usefulness of magnetic resonance volumetric evaluation in predicting response to preoperative con-


