Evaluation of an Objective MRI-Based Tumor Regression Grade (mrTRG) Score and a Subjective Likert Score for Assessing Treatment Response in Locally Advanced Rectal Cancers—A Retrospective Study

Aparna N. Katdare1,2 Akshay D. Baheti1,2• Sayali Y. Pangarkar1,2 Kunal A. Mistry1,2 Suman K. Ankathi1,2 Purvi D. Haria1,2 Amit J. Choudhari1,2 Amrita Guha1,2 Kunal Gala1,2• Nitin Shetty1,2• Suyash Kulkarni1,2 Mukta Ramadwar2,3 Munita Bal2,3

1 Department of Radiodiagnosis, Tata Memorial Hospital, Mumbai, Maharashtra, India
2 Homi Bhabha National Institute, Mumbai, Maharashtra, India
3 Department of Pathology, Tata Memorial Hospital, Mumbai, Maharashtra, India

Address for correspondence Akshay Baheti, MD (Radiodiagnosis), Department of Radiodiagnosis, Tata Memorial Hospital, Mumbai 400012, Maharashtra, India (e-mail: akshaybaheti@gmail.com).

Indian J Radiol Imaging 2024;34:69–75.

Abstract

Purpose: Magnetic resonance imaging (MRI) with the help of MRI-based tumor regression grade (mrTRG) score has been used as a tool to predict pathological tumor regression grade (pTRG) in patients of rectal cancer post-neoadjuvant chemoradiation. Our study aims to evaluate the ability of MRI in assessing treatment response comparing an objective mrTRG score and a subjective Likert score, with a focus on the ability to predict pathologic complete response (pCR).

Methods: Post-treatment MRI studies were retrospectively reviewed for 170 consecutive cases of histopathologically proven rectal cancer after receiving neoadjuvant chemoradiation and prior to surgery by two oncoradiologists blinded to the eventual postoperative histopathology findings. An objective (mrTRG) and a subjective Likert score were assigned to all the cases. Receiver operating characteristic curves were constructed to determine the ability of Likert scale and mrTRG to predict pCR, with postoperative histopathology being the gold standard. The optimal cutoff points on the scale of 1 to 5 were obtained for mrTRG and Likert scale with the greatest sum of sensitivity and specificity using the Youden Index.

Results: The most accurate cutoff point for the mrTRG to predict complete response was 2.5 (using Youden index), with a sensitivity of 69.2%, specificity of 69.6%, positive predictive value (PPV) of 85.6%, negative predictive value (NPV) of 46.4%, and accuracy of 69.3%. The most accurate cutoff for the Likert scale to predict complete response was 3.5, with a sensitivity of 47.5%, specificity of 89.1%, PPV of 91.9%, NPV of 39.4%, and accuracy of 59%. mrTRG had a lower cutoff and was more accurate in predicting pCR compared to Likert score.

Keywords

- Likert
- MRI
- rectal cancer
- TRG

© 2023. Indian Radiological Association. All rights reserved.
This is an open access article published by Thieme under the terms of the Creative Commons Attribution-NonDerivative-NonCommercial-License, permitting copying and reproduction so long as the original work is given appropriate credit. Contents may not be used for commercial purposes, or adapted, remixed, transformed or built upon. (https://creativecommons.org/licenses/by-nc-nd/4.0/)

Thieme Medical and Scientific Publishers Pvt. Ltd., A-12, 2nd Floor, Sector 2, Noida-201301 UP, India
Introduction

Magnetic resonance imaging (MRI) has an established role in baseline staging, post-treatment restaging, and surgical planning in rectal cancer. Preoperative chemoradiation is used to reduce tumor bulk and sterilize the surgical field. Apart from killing tumor cells, it also induces inflammation, fibrosis, submucosal edema, and mucinous change in the tumor. The tumor response to treatment is categorized on histopathology using pathological tumor regression grade (pTRG), which correlates with patient outcomes.2,3

Few MRI studies have suggested that an MRI-based tumor regression grade (mrTRG) on the pre-operative MRI can be used to predict pTRG; however, the reported accuracy of mrTRG varies in different studies.4–7 An objective mrTRG may be unable to account for findings like mucinous changes or thick fibrosis with minimal heterogeneity or irregularity. A subjective Likert scale may be able to assess these better. A study analyzing prostatic lesions and comparing Likert score and presence of mucin (cellular and acellular), was documented in all patients. The method of Mandard et al13 was used to evaluate pTRG. pTRG 1 indicated complete response; pTRG 2 indicated predominant fibrosis with rare residual cancer cells scattered through the fibrosis; pTRG 3 indicated presence of fibrosis and tumor cells, with fibrosis still being predominant; pTRG 4 indicated residual cancer outgrowing fibrosis; and pTRG 5 indicated absence of regressive changes.

Clinical and Histopathology Data

The demographic data, histopathology, and treatment details for the patient were taken from the institutional electronic medical records. The surgical specimens were reported by dedicated gastrointestinal oncopathologists at our institute. The detailed final histopathology, including pTRG, pathological T stage (pT), pathological N stage (pN) and presence of mucin (cellular and acellular), was documented in all patients. The method of Mandard et al13 was used to evaluate pTRG. pTRG 1 indicated complete response; pTRG 2 indicated predominant fibrosis; pTRG 3 indicated presence of fibrosis and tumor cells, with fibrosis still being predominant; pTRG 4 indicated residual cancer outgrowing fibrosis; and pTRG 5 indicated absence of regressive changes.

Imaging and Image Analysis

MRI studies were performed on 1.5T (GE Healthcare, Milwaukee, Wisconsin, United States), 1.5T (Philips Medical Systems, Eindhoven, the Netherlands), or 3T (GE Healthcare, Milwaukee, Wisconsin, United States) machines in our institute using the institutional protocol for MRI rectum. This included large field of view (FOV) T1-weighted (T1W) and T2-weighted (T2W) axial sequences, T2W sagittal sequence of the pelvis, high-resolution small FOV thin oblique axial and oblique coronal T2W sequences along the plane of the rectum, and axial DWI (Supplementary Material – Table 1 in Appendix). Intravenous contrast was not administered in accordance with the institutional protocol.

The images were reviewed on Centricity PACS (GE Healthcare, Milwaukee, Wisconsin, United States) workstation by two oncoradiologists in consensus, with 11 and 9 years of experience. The radiologists were blinded to the histopathology findings. An objective MR predicted tumor regression grade (mrTRG) was determined (Supplementary Material – Table 1).14 A 5-point subjective Likert scale was also assigned with score of 1 to 5, Likert 1 meaning highly likely to be pCR and Likert 5 meaning highly unlikely to be pCR (Supplementary Material). DWI and apparent diffusion coefficient (ADC) images were visually assessed as well. For patients without DWI (n = 15), mrTRG was based purely on T2WI; presence of thick or irregular

Materials and Methods

Subjects

This is a retrospective study approved by the Institutional Review Board. One-hundred seventy consecutive cases of histopathologically proven rectal cancer were identified from the database of our colorectal disease management group over 5 years. We excluded patients operated more than 8 weeks after MRI (1 patient) and MRIs lacking adequate diagnostic quality (3 patients). All the remaining 166 patients had an MRI done after receiving neoadjuvant chemoradiation and prior to surgery. One-hundred forty-one of one-hundred sixty-six patients received presurgical chemoradiation followed by surgery. Twenty-five of one-hundred sixty-six patients that did not respond adequately to the chemoradiation on MRI and clinical examination received a 6-week course of chemotherapy, followed by a second post-treatment MRI and subsequent surgery. All 166 cases underwent definitive surgery within 8 weeks of their latest pre-surgical MRI. We included patients that had a baseline MRI (36/166) or post-treatment MRI (8/166) done outside our institute but of optimum diagnostic quality, with the Digital Imaging and Communications in Medicine (DICOM) images available to be uploaded on the institutional picture archiving and communication system (PACS). We also included patients without a baseline MRI and only the post-treatment MRI on PACS (56/166), since only the post-treatment MRIs were being assessed. Diffusion-weighted imaging (DWI) sequence was available in 151/166 cases.

Conclusion: An objective mrTRG was more accurate than a subjective Likert scale to predict complete response in our study.
Table 1 Summary of MRI regression scale used in our study

<table>
<thead>
<tr>
<th>mrTRG scale</th>
<th>mrTRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Radiological complete response: no evidence of the treated tumor, or presence of T2 hypointense fibrosis without diffusion restriction</td>
</tr>
<tr>
<td>2</td>
<td>Good response (T2 hypointense fibrosis with restricted diffusion or minimal mucinous changes, without definite T2 intermediate signal intensity tumor)</td>
</tr>
<tr>
<td>3</td>
<td>Moderate response (visible T2 intermediate signal signifying tumor with variable fibrosis or mucinous changes)</td>
</tr>
<tr>
<td>4</td>
<td>Slight response (little areas of fibrosis or mucin; predominantly residual tumor)</td>
</tr>
<tr>
<td>5</td>
<td>No response or worsening disease</td>
</tr>
</tbody>
</table>

Abbreviations: MRI, magnetic resonance imaging; mrTRG, MRI-based tumor regression grade.

fibrrosis was scored mrTRG2 while minimal fibrosis was scored mrTRG 1.

Statistical Analysis
The SPSS (the statistical package for social sciences), IBM Corp, released 2017, IBM SPSS Statistics for Windows, Version 25.0. Armonk, NY: IBM Corp was used for statistical analysis. Receiver operating characteristic curves were constructed to determine the ability of Likert scale, mrTRG and DWI to predict complete response, with postoperative histopathology being the gold standard. The optimal cutoff points on the scale of 1 to 5 were obtained for mrTRG and Likert scale with the greatest sum of sensitivity and specificity using the Youden Index. Further, the sensitivity, specificity, positive likelihood ratio, negative likelihood ratio, positive predictive value (PPV), negative predictive value (NPV), and accuracy with 95% confidence intervals (CI) were calculated for mrTRG, Likert scale, and DWI sequence.

Results
One-hundred nine of one-hundred sixty-six (65.7%) patients were males, with a median age of 50 years (range: 29–71 years). Fifty-four of one-hundred sixty-six (33%) patients had pCR on pathology. The mrTRG, Likert, and pTRG distribution of the total number of patients is shown in – Fig. 1.

The respective distribution of pTRG among patients assigned different Likert scores and pTRGs is given in – Table 3. Six of nine cases (67%) that were Likert 1 showed pCR (pTRG 1), 2/9 (22.2%) showed pTRG 2, and 1/9 (11%) was pTRG 3. Out of the 35 cases that were Likert 2, 21/35 (60%) were pCR, while 7/35 (20%) were pTRG 2. If Likert scores 1 and 2 were considered together, then 27/44 (61%) cases were pCR, while 9/44 (20%) cases were pTRG 2. The percentage of pCR for different Likert scores is given in – Table 4.

With regard to mrTRG (– Table 5), 8/9 (89%) cases of mrTRG 1 were pCR, 1/9 was pTRG 2, and none was pTRG 3-5. Among the 60 cases that were mrTRG 2, 29/60 (48%) were pCR, 15/60 (25%) were pTRG 2, and 16/60 (27%) were pTRG 3-5 (– Fig. 2). If mrTRG 1 and 2 were considered together, then 37/69 (54%) cases were pCR and 16/69 (23%) were pTRG 2. The percentage of pCR for different mrTRG is given in – Table 6.

The most accurate cutoff point for the mrTRG to predict complete response was 2.5 (using Youden index), with a sensitivity of 69.2%, specificity of 69.6%, PPV of 85.6%, NPV of 46.4%, and accuracy of 69.3%. The most accurate cutoff for the Likert scale to predict complete response was 3.5, with a sensitivity of 47.5%, specificity of 89.1%, PPV of 91.9%, NPV of 39.4%, and accuracy of 59%. Thus, overall, mrTRG had a lower cutoff and was more accurate in predicting pCR compared to Likert score.

Ninety of one-hundred sixty-six cases showed no mucinous change. 46/166 showed mild mucinous changes, and 30/166 showed moderate to substantial mucinous changes. None of the latter 30 cases were labeled mrTRG 1, with one being labelled Likert 1. Eleven of thirty (36.67%) cases demonstrated complete response on pathology.

One-hundred fifty-one patients had DWI available. Sixty-seven of one-hundred fifty-one (44%) residual tumors demonstrated restricted diffusion, with 55/67 (82%) of these not having pCR. Of the 84/151 cases without restricted diffusion, 32/84 (38%) had pCR. The overall sensitivity of presence of restricted diffusivity on DWI to predict residual disease was 51.4%, with specificity of 72.7%, PPV of 82.1%, NPV of 38.1%, and accuracy of 57.6%.

Discussion
Our study demonstrates that an objective mrTRG to be more accurate than a subjective Likert score in predicting pTRG. Overall, both Likert and mrTRG scores tended to be higher leading to over-staging and under-prediction of complete
response. Also, MRI was not very accurate in predicting pTRG, consistent with other studies. This underlines the fact that mrTRG is not a parameter that can be reliably assessed in the post-neoadjuvant treatment setting, and need not be reported. A clinically more relevant parameter for assessing on MRI is the presence of CR or near complete response, identifying patients for potential “Watch and Wait” strategy. The incidence of pCR in our study was 33%, comparable with other studies that report an incidence of 8 to 38%.

The MERCURY study elaborates the morphologic responses to chemoradiation in the form of fibrosis or desmoplastic reaction and mucinous change. Fibrosis presents as T2 dark spicules or strands radiating from the rectal wall. According to the objective mrTRG criteria, T2 dark signal intensity areas represent fibrosis and intermediate signal intensity areas represent residual tumor. We found

### Table 3 Case-wise correlation of Likert score and pTRG

<table>
<thead>
<tr>
<th>pTRG</th>
<th>Likert</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>6 (67%)</td>
<td>21 (60%)</td>
<td>21</td>
<td>4</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2 (22%)</td>
<td>7 (20%)</td>
<td>18</td>
<td>15</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>1 (11%)</td>
<td>4 (30%)</td>
<td>12</td>
<td>5</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>10</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
<td>35</td>
<td>60</td>
<td>37</td>
<td>25</td>
<td></td>
</tr>
</tbody>
</table>

Abbreviation: pTRG, pathological tumor regression grade.

### Table 4 Number of cases showing pCR when different subsets of Likert score are considered

<table>
<thead>
<tr>
<th>Likert</th>
<th>No. of cases (n)</th>
<th>pCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>6 (67%)</td>
</tr>
<tr>
<td>1 + 2</td>
<td>44</td>
<td>27 (61%)</td>
</tr>
<tr>
<td>3 + 4 + 5</td>
<td>122</td>
<td>28 (23%)</td>
</tr>
</tbody>
</table>

Abbreviation: pCR, pathologic complete response.

### Table 5 Case-wise correlation of mrTRG and pTRG

<table>
<thead>
<tr>
<th>pTRG</th>
<th>mrTRG</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8 (89%)</td>
</tr>
<tr>
<td>2</td>
<td>1 (11%)</td>
</tr>
<tr>
<td>3</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>1</td>
</tr>
<tr>
<td>Total</td>
<td>9</td>
</tr>
</tbody>
</table>

Abbreviations: mrTRG, magnetic resonance imaging-based tumor regression grade; pTRG, pathological tumor regression grade.
these criteria to be more accurate (70%) than the subjective Likert scale with an accuracy of 59%. Our study results showed concordant findings to a meta-analysis by Jang et al., which showed an accuracy of 77 to 82% for MrTRG1 for diagnosis of pathological T1 or lower and an accuracy of 56 to 74% for MrTRG1-2 for diagnosis of pathological T1 or lower. We believe that this can be attributed to T2 shine through effect and ghosting artifact produced by pools of mucin.

The role of DWI in accurately assessing response is not clear. In a meta-analysis of 14 studies by Wu et al for assessing treatment response in rectal cancer, DWI did not improve sensitivity when coupled with conventional MRI. In another meta-analysis of 33 studies by van der Paardt et al, the pooled sensitivity of studies that included DWI in their protocol was higher (83.6 vs. 50.4%) than studies that did not. Our study also found limited sensitivity and accuracy of DWI in detecting residual disease. The basis of DWI lies in differential movement of water molecules in various tissues.

Table 6 Number of cases showing pCR when different subsets of MrTRG are considered

<table>
<thead>
<tr>
<th>MrTRG</th>
<th>No. of cases (n)</th>
<th>pCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>9</td>
<td>8 (89%)</td>
</tr>
<tr>
<td>1 + 2</td>
<td>69</td>
<td>37 (54%)</td>
</tr>
<tr>
<td>3 + 4 + 5</td>
<td>97</td>
<td>18 (19%)</td>
</tr>
</tbody>
</table>

Abbreviations: MrTRG, magnetic resonance imaging-based tumor regression grade; pCR, pathologic complete response.
In an experimental study by Hein et al., comparing the results of DWI in post-treatment setting with histopathology, they found that radiation-induced fibroinflammatory changes can also reduce the free movement of extracellular water molecules, thus decreasing ADC values. This may lead to misinterpretation as residual disease, potentially explaining why we observed pCR in 18% patients demonstrating restricted diffusion.

A subjective Likert scale largely depends on experience of the observer. Though it was comparatively more specific (89%) compared to mrTRG (69%) in our study, it showed a poor sensitivity of 47% and had an overall lower accuracy, making mrTRG a more useful method of assessment for an inexperienced learner.

The moderate accuracy of MRI in predicting treatment response emphasizes the need for additional methods of assessment such as endoscopy and per rectal examination in order to qualify a patient for a nonsurgical Watch and Wait assessment such as endoscopy and per rectal examination in response emphasizes the need for additional methods of making mrTRG a more useful method of assessment for an inexperienced learner.

The study had various limitations such as a small cohort, retrospective analysis, and radiological assessment by only two radiologists without analysis of interobserver variation. Further randomized and prospective studies would be needed to decide on reliable parameters for prediction of pCR.

In conclusion, objective mrTRG was more accurate than a subjective Likert score to predict complete response in our study.

Conflict of Interest
None declared.

References
11 Arya S, Das D, Engineer R, Saklani A. Imaging in rectal cancer with emphasis on local staging with MRI. Indian J Radiol Imaging 2015; 25(02):148–161
15 Voogt ELK, Nordkamp S, van Zogbel DMGJ, et al. MRI tumour regression grade in locally recurrent rectal cancer. BJS Open 2022; 6(03):zrac033