Title

Prospective Correlation of Magnetic Resonance Tumor Regression Grade With Pathologic Outcomes in Total Neoadjuvant Therapy for Rectal Adenocarcinoma.

Journal

Journal of Clinical Oncology, 41(29)

Authors

Hall, William

Li, Jiahe

You, Y

et al.

Publication Date

2023-10-10

DOI

10.1200/JCO.22.02525

Peer reviewed

Prospective Correlation of Magnetic Resonance Tumor Regression Grade With Pathologic Outcomes in Total Neoadjuvant Therapy for Rectal Adenocarcinoma

William A. Hall, MD[1]; Jiahe Li, MS[2]; Y. Nancy You, MD[3]; Marc J. Gollub, MD[4]; Joseph R. Grajo, MD[5,6]; Mark Rosen, MD[7]; Greg dePrisco, MD[8]; Greg Yothers, PhD[2]; Jennifer A. Dorth, MD[9]; Osama E. Rahma, MD[10]; Marcia M. Russell, MD[11]; Howard M. Gross, MD[12]; Samuel A. Jacobs, MD[13]; Bryan A. Faller, MD[14]; Sagila George, MD[15]; Tareq Al baghdadi, MD[16]; Michael G. Haddock, MD[17]; Richard Valicenti[18]; Theodore S. Hong, MD[19]; and Thomas J. George, MD[5,6]

DOI https://doi.org/10.1200/JCO.22.02525

ABSTRACT

PURPOSE Total neoadjuvant therapy (TNT) is a newly established standard treatment for rectal adenocarcinoma. Current methods to communicate magnitudes of regression during TNT are subjective and imprecise. Magnetic resonance tumor regression grade (MR-TRG) is an existing, but rarely used, regression grading system. Prospective validation of MR-TRG correlation with pathologic response in patients undergoing TNT is lacking. Utility of adding diffusion-weighted imaging to MR-TRG is also unknown.

METHODS We conducted a multi-institutional prospective imaging substudy within NRG-GI002 (ClinicalTrials.gov identifier: NCT02921256) examining the ability of MR-based imaging to predict pathologic complete response (pCR) and correlate MR-TRG with the pathologic neoadjuvant response score (NAR). Serial MRIs were needed from 110 patients. Three radiologists independently, then collectively, reviewed each MRI for complete response (mriCR), which was tested for positive predictive value (PPV), negative predictive value (NPV), sensitivity, and specificity with pCR. MR-TRG was examined for association with the pathologic NAR score. All team members were blinded to pathologic data.

RESULTS A total of 121 patients from 71 institutions met criteria: 28% were female (n = 34), 84% White (n = 101), and median age was 55 (24-78 years). Kappa scores for T- and N-stage after TNT were 0.38 and 0.88, reflecting fair agreement and near-perfect agreement, respectively. Calling an mriCR resulted in a kappa score of 0.82 after chemotherapy and 0.56 after TNT reflected near-perfect agreement and moderate agreement, respectively. MR-TRG scores were associated with pCR ($P < .01$) and NAR ($P < .0001$), PPV for pCR was 40% (95% CI, 26 to 53), and NPV was 84% (95% CI, 75 to 94).

CONCLUSION MRI alone is a poor tool to distinguish pCR in rectal adenocarcinoma undergoing TNT. However, the MR-TRG score presents a now validated method, correlated with pathologic NAR, which can objectively measure regression magnitude during TNT.

ACCOMPANYING CONTENT

Appendix

Data Sharing Statement

Data Supplement

Protocol

Accepted May 9, 2023

Published July 21, 2023

J Clin Oncol 41:4643-4651

View Online Article

INTRODUCTION

Contemporary management of locally advanced rectal adenocarcinoma now includes total neoadjuvant therapy (TNT).[1-3] The TNT strategy includes both systemic chemotherapy as well as various types of radiotherapy (RT), all given before surgical resection.[3,4] Optimal sequencing and dosage of both chemotherapy and RT remains an area of active investigation.[5,6] There are multiple reasons for the widespread adoption of TNT in rectal cancer management, including improved preoperative chemotherapy tolerance, increased response rates, higher probability of organ preservation, and the ability to assess tumor response to therapy.[6,7] Studies demonstrate that 30%-50% of patients can achieve nonoperative management when undergoing TNT.[3] This creates a considerable need for accurate clinical characterization of tumor regression. Accurate and objective characterization of regression would enable clear communication of response magnitude across oncologic teams, allowing imaging to

CONTEXT

Key Objective

Characterization of rectal tumor regression during total neoadjuvant therapy (TNT) is a critically important task. Currently, minimal data exist on how reliably rectal MRI can characterize regression during TNT. We prospectively validated the ability of a numerical grading scale to predict pathologic complete response (pCR) and correlate with tumor regression.

Knowledge Generated

The magnetic resonance tumor regression grade (MR-TRG) is a poor tool to predict pCR to TNT. However, MR-TRG does correlate with the magnitudes of pathologic regression during TNT. Finally, diffusion-weighted imaging improves the AUC when compared with MR-TRG alone.

Relevance *(A.H. Ko)*

This study highlights both the usefulness and limitations of MRI in predicting pathologic tumor response during TNT for rectal adenocarcinoma. While MR-TRG correlates with the magnitude of pathologic regression, it is not an accurate predictor of complete pathologic response.*

*Relevance section written by *JCO* Associate Editor Andrew H. Ko, MD, FASCO.

serve as an integral biomarker, used to personalize durations of neoadjuvant therapy.

Over the past decade, MRI has emerged as the standard of care for staging patients with rectal adenocarcinoma.[8] When patients undergo TNT, changes on rectal MRI are often used to characterize response. Despite this, current strategies to communicate magnitudes of response are vague or subjective. Correlation between existing strategies and magnitudes of pathologic response is unclear. Establishing a reliable tumor response classification to TNT represents an important area of research in rectal adenocarcinoma with currently no prospectively validated and objective imaging scores available.[9]

The magnetic resonance tumor regression grade (MR-TRG) presents the magnitude of tumor regression as an ordinal score, ranging from 1 to 5.[10] MR-TRG has been studied in the non-TNT setting with multiple retrospective, and small prospective, series.[11-16] Most of these studies report MRIs acquired in single institutions, resulting in limitations to the generalizability of conclusions.[17] Additionally, the benefit of more complex diffusion-weighted imaging (DWI) as part of MR-based assessment is unknown. To our knowledge, no studies examining MR-TRG have included multi-institutional cohorts prospectively obtained from patients undergoing TNT.

To comprehensively address these knowledge gaps in rectal MRI, we conducted a prospective imaging study within the US National Clinical Trials Network (NCTN) via the NRG-GI002 trial (ClinicalTrials.gov identifier: NCT02921256). GI002 is a National Cancer Institute (NCI)–sponsored phase II platform study examining novel radiosensitizers for rectal adenocarcinoma using TNT for all patients, with a primary end point of pathologic regression, measured by the pathologic neoadjuvant response (NAR) score.[18,19] The NAR score is a validated pathologic grading system that describes the magnitude of tumor regression seen when comparing the initial clinical stage to the pathologic stage after neoadjuvant therapy.[18] For this imaging substudy, the primary end point is to characterize the reliability of MRI to assess pathologic complete response (pCR). Additionally, we sought to investigate for a correlation between the MR-TRG and pathologic NAR.[18,20,21] We hypothesized the MR-TRG score would objectively characterize a pCR and correlate with NAR.

METHODS

Patients

Patients for this study were enrolled on the NRG-GI002 prospective clinical trial with previously reported eligibility criteria and clinical end points.[20,21] NRG-GI002 was approved by local human investigations committees/institutional review boards in accordance with assurances approved by the Department of Health and Human Services. Written informed consent was required. Briefly, patients had locally advanced rectal cancer, defined as distal location (cT3-4 ≤ 5 cm from anal verge, any N); bulky (any cT4 or tumor within 3 mm of mesorectal fascia); high risk for metastatic disease (cN2); or not a sphincter-sparing surgery candidate. GI002 included two sequential experimental arms plus a concurrent control arm. All patients received 4 months of FOLFOX systemic chemotherapy followed by long-course RT concurrent with capecitabine, followed by surgical resection. The first experimental arm included the addition of veliparib (a poly [ADP-ribose] polymerase inhibitor), concurrent with RT.[20] The second

experimental arm tested addition of pembrolizumab (an antibody directed against the PD-1 receptor), concurrent with and after RT.[21] The primary end point of the parent study was pathologic NAR score, which is calculated using the following formula: NAR = [5pN − 3(cT − pT) + 12]²/9.61.[20] Patients staged with MRI were eligible for the imaging biomarker substudy, which was an a priori designed aspect of the parent trial.

Imaging Substudy Design and Inclusion Criteria

This imaging companion study was designed to characterize the ability of MRI to serve as a surrogate for pCR, using the MR-TRG scoring system. In other words, an MR-TRG of 1 was considered a radiologic description of pCR. We also sought to understand the utility (if any) to using DWI, which was also incorporated into response assessment and was compared with MR-TRG. Three radiologists generated independent binary assessments of the presence of a possible pCR and were blinded regarding pathologic surgical results.

Assuming a one-sided alpha = 10%, with 80% power, and a null hypothesis that the positive predictive value (PPV) of radiologic MR-TRG for pCR is 80%, a total of n = 22 positive calls (ie, 22 complete response events described on MRI) were necessary to reject the null if the true PPV were 95%. If approximately 20% of treated patients resulted in a pCR, which was expected on the basis of the existing literature, then 110 unique patients were necessary with both pre- and post-TNT MRIs.

Scoring Criteria, MRI Acquisition, and Assessment

The MRI review process consisted of two different standardized forms (baseline and post-treatment) used by each independent central radiologist. Diagnostic radiologists, from different institutions, reviewed scans and completed forms independently. Then they met in person and/or virtually, to complete a consensus read. Radiology forms were designed before study initiation and remained unchanged (Data Supplement, online only). Specific MRI protocols were recommended in the parent trial and are included in the trial supplement. Main parameters included radiologic T- and N-stage, along with nodal measurement grouping when nodes were between 5-9 mm and ≥10 mm. Number of nodes with heterogeneous signal, rounded shape, and irregular margins were recorded. Presence of extramural vascular invasion was also recorded.

Post-treatment MR-TRG was scored ranging from 1 to 5, which was identical to the scoring criteria used in prior publications of MR-TRG. MR-TRG 1 indicated no/minimal fibrosis visible (tiny linear scar) and no tumor signal; MR-TRG 2 indicated dense fibrotic scar (low-signal intensity) but no macroscopic tumor signal (indicates no or microscopic tumor present); MR-TRG 3 indicated fibrosis predominates, but obvious measurable areas of tumor signal present; MR-TRG 4 indicated tumor signal predominates with little or minimal fibrosis; and MR-TRG 5 indicated the considered tumor-signal only, no fibrosis present, and included tumor progression.[10] When available, DWI data were also examined for each MRI. High signal intensity on high b-value images corresponded to low signal intensity on the apparent diffusion coefficient map, aka restricted diffusion. DWI scoring was considered either present (residual tumor likely exists), absent (probable complete response), or equivocal (poor quality or artifact). The final impression for radiologic complete response (mriCR) was assessed as 1 (not present), or alternatively, if the mriCR was present, it was either score as a 2 (present on both the MR-TRG score and DWI), 3 (MR-TRG only), or 4 (DWI only). Full details are provided in the scoring forms (Data Supplement). Figure 1 displays examples of MR-TRG response.

Study Support and Image Review/Management

This imaging study was funded by the NCI through the Biomarkers, Imaging, and Quality of Life Studies Funding Program. Images were centrally stored by the Imaging and Radiation Oncology Core in the Transmission of Imaging and Data system, with standardized reports in Medidata Rave. After independent readings, consensus adjudication was performed for cases of read discordance (49%) during in-person and/or virtual meetings (n = 121 cases) with consensus achieved on 117 on the basis of discussion and agreement. Very rarely when consensus among radiologists could not be achieved, mostly because of image quality, the case was excluded from consensus reporting (n = 4).

Exploratory Modifications in the MR-TRG

Additional permutations of the MR-based tumor regression modeling were examined. First was a radiologic NAR (rNAR) score, which was intended to serve as a surrogate for and replacement of the pathologic NAR score for future clinical trials, because pathologic NAR is not measurable in patients for whom surgery is omitted (for any reason). Modeled after the NAR score, the rNAR was calculated using only pre- and post-treatment clinical stage. A second score was a modification to the TRG score, which accounted for the presence of DWI signal reflecting residual tumor, named the modified MR-TRG (DWImodMR-TRG).[22] If the DWI signal was present (DWI = 1, meaning residual signal was present, indicating tumor), then the MR-TRG–reported score for that patient was increased (DWImodMR-TRG) and was reported as MR-TRG+1. If DWI = 2 (absent), the DWImodMR-TRG was a reduction of the originally reported MR-TRG score, MR-TRG-1. The modified score was equal to MR-TRG when DWI = 3 (equivocal or missing). This DWImodMR-TRG takes integer values from 0 to 6, with 0 to 2 mapping to tumor absent and 3 to 6 mapping to present, expanding the original 1-5 scale for MR-TRG.[22]

Statistical Analysis

Wilcoxon-Mann-Whitney test was used for association of MR-TRG with pCR. Radiologists' binary assessment of

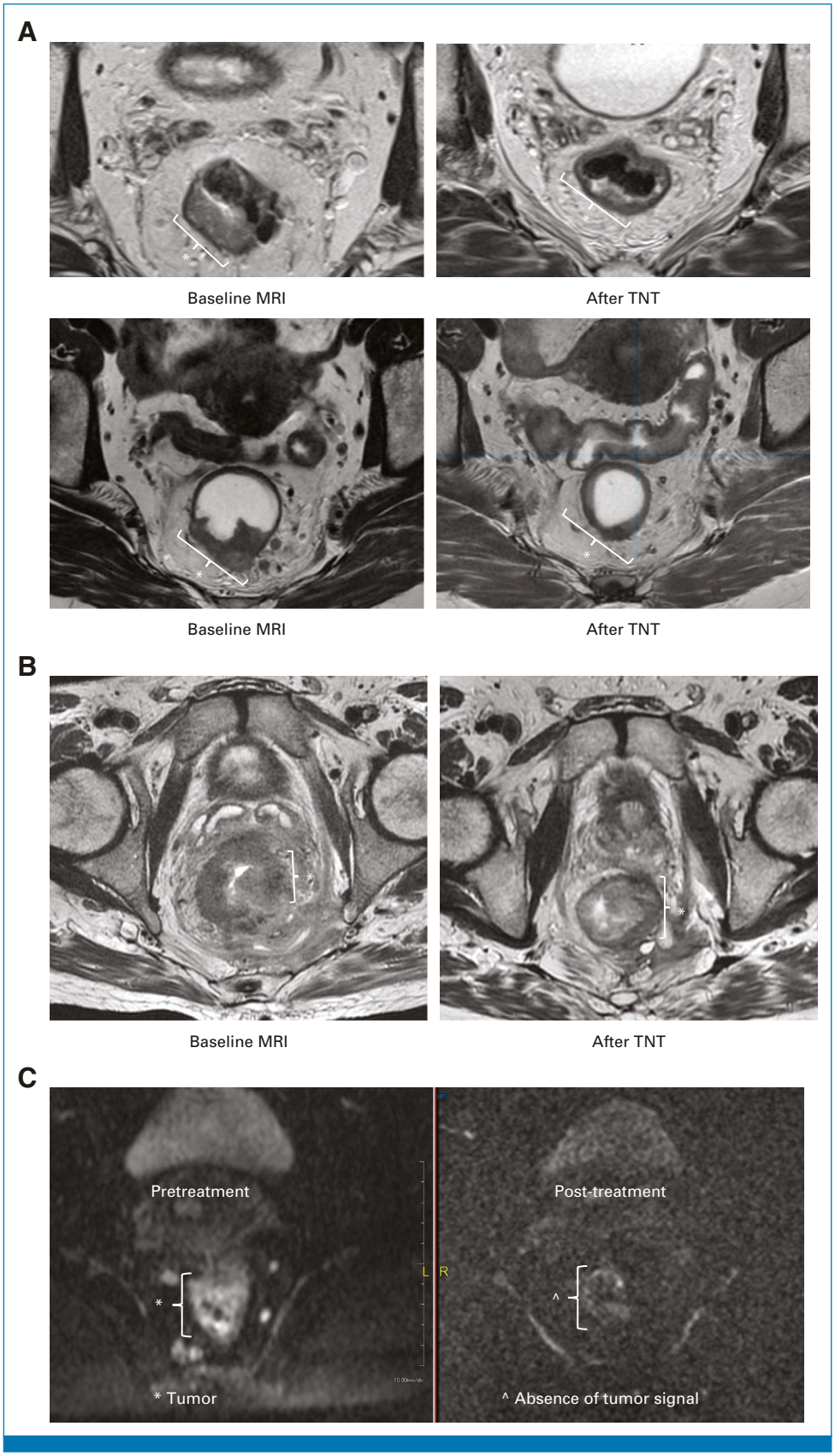

FIG 1. (A) MR-TRG grade 1 example, no tumor evidence after TNT. (B) MR-TRG grade 3 example, residual tumor after TNT. (C) Absence of tumor signal on DWI. DWI, diffusion-weighted imaging; MR-TRG, magnetic resonance tumor regression grade; TNT, total neoadjuvant therapy.

complete response (mriCR) was tested for PPV, negative predictive value (NPV), sensitivity, and specificity with the presence of pCR. Chi-squared test was used for association of mriCR and pCR. Spearman correlation coefficient was used for association of MR-TRG with NAR scores as well as mriCR with NAR. Each diagnostic radiologists' (readers') T-stage, N-stage, MR-TRG score, and DWI assessment was evaluated for inter-reader variability by a generalized linear mixed model-based kappa statistics in the case of ordinal responses and multiple raters.[23,24] Kappa values of agreement were interpreted as follows: poor = <0.0; slight = 0.0-0.2; fair = 0.2-0.4; moderate = 0.4-0.6; substantial = 0.6-0.8; and almost perfect = 0.8-1.0. Receiver operating characteristics (ROC) curves were used to compare MR-TRG, rNAR, and DWImodMR-TRG, with the gold standard, pCR. MR-TRG, mriCR, and DWImodMR-TRG were dichotomized as tumor either present or absent, to conduct comparisons on PPV, NPV, specificity, and sensitivity. Specifically, if DWImodMR-TRG was in 0-2, this was coded as 1 (tumor not present). If it was 3-6, it was coded as 0 (tumor present). Mann-Whitney *U* test was used to compare AUC values. All analyses were performed using SAS 9.4 (SAS Institute, Cary, NC). *P* values < .05 (two-sided) were considered significant.

RESULTS

A total of 121 patients with at least two MRIs centrally available (baseline and post-TNT) accrued to the imaging sub-study; 84 had MRIs at three time points: baseline, post-chemotherapy, and post-chemoRT (post-TNT). These 121 patients were enrolled from 71 different institutions, and three diagnostic radiologists completed individual MRI interpretations across this data set. Additionally, the radiologists completed a consensus read. Of 121 patients, 28% were female (n = 34), 84% White (n = 101), and median age was 55 (24-78 years). Table 1 lists detailed demographic and tumor characteristics.

Correlation of Imaging With pCR and Magnitude of Pathologic Response

Regarding the primary aim of the study, both the individual and consensus radiologist MR-TRG scores were associated with the presence of pCR ($P < .01$) and the pathologic NAR score ($P < .0001$). The correlation between consensus MR-TRG and NAR is 0.43. Dichotomized MR-TRG-alone, mriCR, and DWImodMR-TRG were examined in comparison with the presence of an actual pCR. Table 2 summarizes the consensus of the three radiologists in terms of sensitivity, specificity, PPV, and NPV for cases in which all three scores are available. The highest value measured was NPV at 84% (95% CI, 0.75 to 0.94). mriCR improved specificity over MR-TRG-alone from 0.62 to 0.71 ($P = .07$). Table 3 summarizes specifics of radiologists' response criteria. The Data Supplement includes detailed MR sequences recommended to sites, and Appendix Table A1 (online only) includes the MR-TRG consensus scores compared with pCR categories.

TABLE 1. Patient and Tumor Characteristics for All Randomly Assigned Patients With Paired MRIs: NRG-GI002 Imaging Substudy

	mriCR Not Present		mriCR Present		Total	
Patient or Tumor Characteristic	No.	%	No.	%	No.	%
Sex						
Female	22	30.1	12	25.0	34	28.1
Male	51	69.9	36	75.0	87	71.9
Age, years						
<50	24	32.9	16	33.3	40	33.1
50-59	22	30.1	19	39.6	41	33.9
60-69	23	31.5	11	22.9	34	28.1
≥70	4	5.5	2	4.2	6	5.0
Race						
American Indian or Alaska Native	3	4.1	0	0.0	3	2.5
Asian	4	5.5	3	6.3	7	5.8
Black or African American	4	5.5	1	2.1	5	4.1
Not reported	2	2.7	0	0.0	2	1.7
Unknown	0	0.0	3	6.3	3	2.5
White	60	82.2	41	85.4	101	83.5
Ethnicity						
Hispanic or Latino	1	1.4	1	2.1	2	1.7
Not Hispanic or Latino	72	98.6	45	93.8	117	96.7
Unknown	0	0.0	2	4.2	2	1.7
Distal location						
No	30	41.1	14	29.2	44	36.4
Yes	43	58.9	34	70.8	77	63.6
Bulky						
No	26	35.6	20	41.7	46	38.0
Yes	47	64.4	28	58.3	75	62.0
High risk for metastatic disease						
No	28	38.4	34	70.8	62	51.2
Yes	45	61.6	14	29.2	59	48.8
Not a candidate for SSS						
No	38	52.1	26	54.2	64	52.9
Yes	35	47.9	22	45.8	57	47.1
N stage						
N0	11	15.1	14	29.2	25	20.7
N1	17	23.3	20	41.7	37	30.6
N2	45	61.6	14	29.2	59	48.8
T stage						
T1/T2	1	1.4	1	2.1	2	1.7
T3	45	61.6	39	81.3	84	69.4
T4	27	37.0	8	16.7	35	28.9
pCR						
No	56	80.0	22	57.9	78	72.2
Yes	14	20.0	16	42.1	30	27.8
Total	73	100.0	48	100.0	121	100.0

Abbreviations: mriCR, MRI complete response; pCR, pathologic complete response; SSS, sphincter-sparing surgery.

TABLE 2. Characterization of Consensus Response Scoring Compared With Pathologic Complete Response: NRG-GI002 Substudy

Imaging End Point	Summary
MR-TRG alone[a]	
Sensitivity	19/28 = 68% (95% CI, 0.51 to 0.85)
Specificity	48/77 = 62% (95% CI, 0.52 to 0.73)
Positive predictive value	19/48 = 40% (95% CI, 0.26 to 0.53)
Negative predictive value	48/57 = 84% (95% CI, 0.75 to 0.94)
mriCR[a]	
Sensitivity	16/28 = 57% (95% CI, 0.39 to 0.75)
Specificity	55/77 = 71% (95% CI, 0.61 to 0.82)
Positive predictive value	16/38 = 42% (95% CI, 0.26 to 0.58)
Negative predictive value	55/67 = 82% (95% CI, 0.73 to 0.91)
DWImodMR-TRG[a] (when available[b])	
Sensitivity	18/28 = 64% (95% CI, 0.47 to 0.82)
Specificity	49/77 = 64% (95% CI, 0.53 to 0.74)
Positive predictive value	18/46 = 39% (95% CI, 0.25 to 0.53)
Negative predictive value	49/59 = 83% (95% CI, 0.73 to 0.93)

Abbreviations: DWI, diffusion-weighted imaging; MR-TRG, magnetic resonance tumor regression grade.
[a]MR-TRG, mriCR, and DWImodMR-TRG were dichotomized as tumor present or absent to conduct further comparisons, and 105 patients were included secondary to the elimination of mucinous tumors from inclusion in these final metrics.
[b]DWI was available/interpretable in 59% of patients.

Interobserver Variability Between Central-Central Readers and Central-Local Readers

Interobserver variability was assessed across the reports for the T-stage, N-stage, and MR-TRG scores, along with binary impression of complete response after TNT for each of the central readers. Kappa statistics for the central-central comparison on T-stage was 0.38 (95% CI, 0.34 to 0.42) reflecting fair agreement; for the N-stage was 0.88 (95% CI, NA) reflecting near-perfect agreement; for the interpretation of the MR-TRG was 0.18 (95% CI, 0.12 to 0.23) reflecting slight agreement; and was 0.29 (95% CI, 0.24 to 0.33) for the DWI findings, reflecting fair agreement. Regarding the final impression of complete response (ie, mriCR), the kappa statistic was 0.82 (95% CI, NA), implying the three central reviewers had almost perfect independent agreement after chemotherapy; however, this decreased to 0.56 after TNT. This final agreement was present before the consensus session took place and was not influenced by the time period of the study over which the final impression was made. When comparing the T- and N-stage between the central and local readers, the kappa statistics were 0.34, reflecting fair agreement on T-stage, and 0.59, reflecting moderate agreement with reported N-stage. Interobserver variability did not change with the time of the reports over the years during which the study was conducted for neither the T- nor the N-stage, indicating that interdigitated consensus meetings did not appear to influence the agreement.

Performance of Additional MR-TRG Variations

The rNAR correlated with the pathologic NAR (r = 0.42, P = .0001) and pCR (P = .03). DWImodMR-TRG was also correlated with the NAR score (r = 0.47, P < .0001) and pCR (P = .0001). Figure 2 displays the ROC curves associated with these comparisons. The DWImodMR-TRG trended toward outperformance of the MR-TRG alone, with AUC values of 0.75 (95% CI, 0.62 to 0.87) and 0.70 (95% CI, 0.57 to 0.83) for

TABLE 3. MR-TRG Response Criteria and DWI Scoring Criteria: NRG-GI002 Substudy

Imaging End Point	Definition	Patients in Associated Categories			
DWI	Definition	Missing	pCR-0	pCR-1	Total
1—Present	Residual tumor likely exists	3	38	5	46
2—Absent	Probable clinical complete response	6	9	10	25
3—Equivocal	Quality poor or artifact or signal is seen throughout bowel wall and unclear if extra signal in tumor bed	2	14	11	27
Missing		2	17	4	23
Total		13	78	30	121
MR-TRG	Definition	Missing	pCR-0	pCR-1	Total
1	No/minimal fibrosis visible (tiny linear scar) and no tumor signal	4	2	5	11
2	Dense fibrotic scar (low signal intensity) but no macroscopic tumor signal *(indicates no or microscopic tumor)*	5	27	14	46
3	Fibrosis predominates but obvious measurable areas of tumor signal visible	2	35	8	45
4	Tumor signal predominates with little/minimal fibrosis	1	13	1	15
5	Tumor signal only: no fibrosis, includes progression of tumor	0	0	0	0
Missing		1	1	2	4
Total		13	78	30	121

NOTE. DWI—high signal intensity on high b-value image with corresponding low signal on ADC map = restricted diffusion.
Abbreviations: ADC, apparent diffusion coefficient; DWI, diffusion-weighted imaging; MR-TRG, magnetic resonance tumor regression grade; pCR, pathologic complete response.

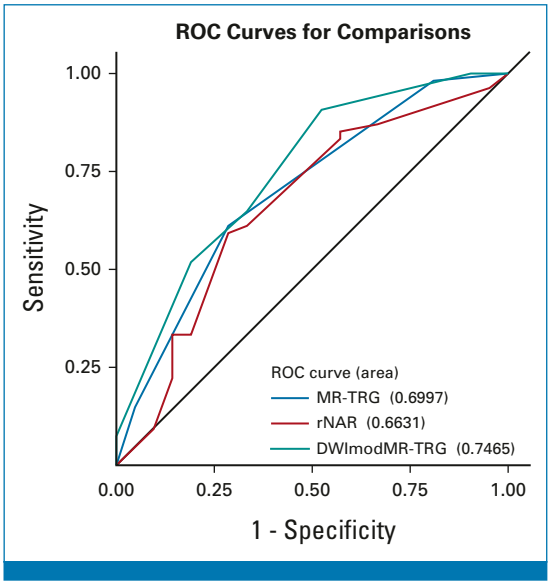

FIG 2. ROC curve comparing MR-TRG with other scoring systems. DWI, diffusion-weighted imaging; MR-TRG, magnetic resonance tumor regression grade; rNAR, radiologic neoadjuvant rectal score; ROC, receiver operating characteristics.

the DWImodMR-TRG and MR-TRG, respectively ($P = .12$). Of note, 23 patients (19%) had missing DWI data and 27 (22%) had equivocal DWI images, which radiologists felt could not be interpreted.

DISCUSSION

An understanding of rectal MRI's ability to accurately describe pCR is of the utmost importance in the era of TNT and rectal cancer organ preservation.[25,26] We have presented a prospective characterization of the MR-TRG ability to accurately describe the presence of pCR across an NCTN trial. The results clearly demonstrate that MRI alone has a poor capability to describe the pCR event. Such results are essential for oncologists to consider when managing patients with rectal cancer treated with TNT.

The PPV of MRI in describing pCR in our study is approximately 40%, no better than a coin flip. This occurred despite three highly specialized radiologists centrally reviewing and agreeing on the findings. Such data clearly identify the limitations of MR-based imaging alone to supplant pathologic assessment and demonstrates the paramount necessity for incorporation of additional clinical metrics of response beyond MRI alone. To our knowledge, this is the first study to present such a characterization of MRIs across a prospective multi-institutional patient cohort being managed with TNT in a uniform manner with consistent treatment, imaging, and pathology acquisition time points. Equally important to consider is the NPV, which demonstrated that between 80% and 90% of the time, MRI correctly identified residual tumor. Such data are imperative for counseling patients regarding the need for surgery.

A second important outcome of this study is the prospective validation of MR-TRG as a strategy to objectively describe rectal tumor regression during TNT. This effort was motivated by the absence of current methods to describe magnitudes of MRI response during TNT. Originally developed nearly a decade ago from the MERCURY trial,[10] the MR-TRG score presents an opportunity to characterize the magnitude of tumor regression objectively. Although follow-up studies have been published using the MR-TRG, adoption of the scoring system within the oncologic community has not taken place. This is likely arising from the retrospective and single-institutional cohort data of existing validation studies, along with the absence of validation in a TNT population. Our data clearly demonstrate that MR-TRG correlates with magnitudes of pathologic regression. Quantifying regression during TNT represents an important objective, because the degree of pathologic regression is well known to be correlated with overall survival.[18,27] With these data, MR-TRG offers a strategy to serve as a quantifiable, objective, and succinct clinical scoring system for use in adaptive prospective TNT trials.

There are also several noteworthy secondary findings from this study. One relates to the modifications of the MR-TRG tested. Our results appear to demonstrate that DWImodMR-TRG trended toward improved AUC values compared with MR-TRG alone; however, this did not reach statistical significance on comparison. Adding DWI data into the MR-TRG score has been reported in a retrospective cohort with an AUC at very impressive levels as high as 0.88.[22] Our data raised the AUC to 0.75 from the consensus MR-TRG alone AUC of 0.70. These findings likely reflect values closer to true accuracy, given the multi-institutional and blinded nature of this study. In previously reported single-institution settings, radiologists' familiarity with the MRI acquisition techniques was high. Notably, DWI was only available in approximately 50% of cases, highlighting the need for the oncologic imaging community to focus on accurate acquisition of such data. Another interesting secondary finding was that after chemotherapy alone, the three central reviewers had almost perfect agreement; however, this decreased substantially after TNT. Etiology of this is unclear and may be related to the difficulty of post-RT interpretation of MR-TRG.

There are limitations to the design of our study. First, pCR is perhaps the utmost conservative metric for nonoperative management and it may be overly conservative. We did not have central pathology review of pCR events. Moreover, our study did not enable prediction of nonoperative management, which is perhaps the most relevant question. Subsequently, expansion of these data is planned for the next NCTN trial, which is focused on organ preservation in rectal cancer. Indeed, pCR is by far the most rigorous outcome we could have attempted to predict using MRI alone, which may have contributed to the modest results. Finally, we are not able to correlate MR-TRG with progression-free or overall

survival, because the data are immature. Continued follow-up of these data is needed (and ongoing) to directly correlate the MR-TRG with overall survival in this population.

In conclusion, our results clearly establish that MR-based determinations of pCR are limited. Such information is critical in a TNT era in which exquisite focus is being placed on selecting patients for nonoperative management. Physical examination and endoscopy are clearly indispensable tools for this task. These findings present compelling data that should provide justification for using the MR-TRG to objectively describe regression more uniformly in patients with rectal adenocarcinoma undergoing TNT. Larger magnitudes of MR-TRG regression correlate with larger magnitudes of pathologic regression, and vice versa. Finally, the NPV of MRI in our study agrees with prior literature and is a helpful, reproducible strength of MRI. When patients do not have evidence of complete response on imaging ~80% of the time, this is confirmed pathologically. Such data can be helpful in counseling patients about the necessity of surgery under such circumstances. This work is ongoing as part of the translational bioimaging correlates of the NRG-GI002 bioimaging study portfolio. Future directions using MR-TRG are planned by incorporation of novel clinical parameters of response assessment, including patient characteristics, tumor characteristics, circulating tumor DNA, endoscopic image evaluation, and MR-based radiomic metrics of response. The generation of a robust clinical/radiomic signature incorporating multiple factors, including DWI, circulating tumor DNA, and clinical regression assessment, is needed and underway.

AFFILIATIONS

[1]Froedtert and the Medical College of Wisconsin, Milwaukee, WI
[2]The University of Pittsburgh, Pittsburgh, PA
[3]University of Texas MD Anderson Cancer Center, Houston, TX
[4]Memorial Sloan Kettering Cancer Center, New York, NY
[5]University of Florida, Gainesville, FL
[6]University of Florida Health Cancer Center, Gainesville, FL
[7]Imaging and Radiation Oncology Core (IROC) Group, and the University of Pennsylvania, Philadelphia, PA
[8]Baylor Scott and White Health Baylor University Medical Center at Dallas, Dallas, TX
[9]University Hospitals Seidman Cancer Center and Case Western Reserve University, Cleveland, OH
[10]Dana-Farber/Harvard Cancer Institute, Boston, MA
[11]Department of Surgery, David Geffen School of Medicine at UCLA, and VA Greater Los Angeles Healthcare System, Los Angeles, CA
[12]Dayton Clinical Oncology Program, Dayton, OH
[13]NSABP Foundation, Pittsburgh, PA
[14]Missouri Baptist Medical Center/Heartland NCORP, St Louis, MO
[15]Stephenson Cancer Center University of Oklahoma Health Sciences Center, Oklahoma City, OK
[16]Trinity Health Ann Arbor Hospital, Michigan Cancer Research Consortium (NCORP), Ann Arbor, MI
[17]Mayo Clinic, Rochester, MN
[18]University of California Davis Comprehensive Cancer Center/UC Davis School of Med/UC Davis Health, Sacramento, CA
[19]Massachusetts General Hospital Cancer Center, Harvard Medical School, Boston, MA

CORRESPONDING AUTHOR

William A. Hall, MD, Department of Radiation Oncology and Department of Surgery, Graduate School of Biomedical Sciences, Froedtert and the Medical College of Wisconsin, Milwaukee, WI; Twitter: @whallradonc; e-mail: whall@mcw.edu.

DISCLAIMER

W.A.H. (Froedtert and the Medical College of Wisconsin, Milwaukee, WI), J.L., MS (The University of Pittsburgh, Pittsburgh, PA), G.Y., PhD (The University of Pittsburgh, Pittsburgh, PA), and T.J.G., MD (University of Florida Health Cancer Center, Gainesville, FL), had full access to all data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis, as well as the work as a whole, from inception to published article. The funders had no role in the design of the study; the collection, analysis, and/or interpretation of the data; the writing of the manuscript; or the decision to submit the manuscript for publication.

PRIOR PRESENTATION

Presented in part at the 2016 ASCO Annual Meeting, Chicago, IL, June 3-7, 2016; 2017 ASCO GI Symposium, San Francisco, CA, January 19-21, 2017; 2017 ASCO Annual Meeting, Chicago, IL, June 2-6, 2017; 2018 ASCO GI Symposium, San Francisco, CA, January 18-20, 2018; 2019 ASCO GI Symposium, San Francisco, CA, January 17-19, 2019; 2019 ASCO Annual Meeting, Chicago, IL, May 31-June 1, 2019; 2021 ASCO GI Symposium, virtual, January 15-17, 2021; and the 2021 American Society for Therapeutic Radiology and Oncology Annual Meeting, Chicago, IL, October 24-27, 2021.

SUPPORT

Supported by U10CA180868, -180822; UG1-189867; U24-196067; 5U24CA180803; BIQSPF grant; AbbVie; and Merck.

CLINICAL TRIAL INFORMATION

NCT02921256

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Disclosures provided by the authors are available with this article at DOI https://doi.org/10.1200/JCO.22.02525.

DATA SHARING STATEMENT

The study Protocol (online only) and informed consent form will be made available. Individual participant data that underlie the results reported in this article, after deidentification, will be available within 1 year after publication and will be accessible through the NCTN Data Archive. Data will be available to researchers who wish to analyze the data in secondary studies to enhance the public health benefit of the original work. Requirements may include (but not be limited to) a research plan, a Data Use Agreement (DUA), and legally binding signatures. (https://nctn-data-archive.nci.nih.gov/about-us#:~:text=%20To%20request%20data%2C%20the%20following%20are%20required%3A,Data%20Use%20Agreement%20%28DUA%29%20containing%20auto-...%20More%20).

AUTHOR CONTRIBUTIONS

Conception and design: William A. Hall, Marc J. Gollub, Joseph R. Grajo, Mark Rosen, Greg dePrisco, Greg Yothers, Marcia M. Russell, Samuel A. Jacobs, Richard Valicenti, Theodore S. Hong, Thomas J. George
Administrative support: Greg Yothers, Richard Valicenti
Provision of study materials or patients: Howard M. Gross, Bryan A. Faller, Tareq Al baghdadi, Richard Valicenti
Collection and assembly of data: Y. Nancy You, Marc J. Gollub, Joseph R. Grajo, Mark Rosen, Greg dePrisco, Greg Yothers, Howard M. Gross, Bryan A. Faller, Tareq Al baghdadi, Richard Valicenti
Data analysis and interpretation: Jiahe Li, Marc J. Gollub, Joseph R. Grajo, Greg dePrisco, Greg Yothers, Jennifer A. Dorth, Osama E. Rahma, Marcia M. Russell, Samuel A. Jacobs, Bryan A. Faller, Tareq Al baghdadi, Michael G. Haddock, Richard Valicenti, Theodore S. Hong, Thomas J. George
Manuscript writing: All authors
Final approval of manuscript: All authors
Accountable for all aspects of the work: All authors

ACKNOWLEDGMENT

The authors acknowledge the contributions of Christine I. Rudock, Publications and Graphics Specialist, and Wendy L. Rea, BA, Editorial Associate, both of whom are employees of NSABP. They were not compensated beyond their normal salaries for this work.

REFERENCES

1. Cercek A, Roxburgh CSD, Strombom P, et al: Adoption of total neoadjuvant therapy for locally advanced rectal cancer. JAMA Oncol 4:e180071, 2018
2. Kasi A, Abbasi S, Handa S, et al: Total neoadjuvant therapy vs standard therapy in locally advanced rectal cancer: A systematic review and meta-analysis. JAMA Netw Open 3:e2030097, 2020
3. Garcia-Aguilar J, Patil S, Gollub MJ, et al: Organ preservation in patients with rectal adenocarcinoma treated with total neoadjuvant therapy. J Clin Oncol 40:2546-2556, 2022
4. Conroy T, Bosset JF, Etienne PL, et al: Neoadjuvant chemotherapy with FOLFIRINOX and preoperative chemoradiotherapy for patients with locally advanced rectal cancer (UNICANCER-PRODIGE 23): A multicentre, randomised, open-label, phase 3 trial. Lancet Oncol 22:702-715, 2021
5. Shi DD, Mamon HJ: Playing with dynamite? A cautious assessment of TNT. J Clin Oncol 39:103-106, 2021
6. Cercek A, Romesser PB, Crane CH, et al: In defense of TNT: A dynamite strategy. J Clin Oncol 39:1185-1186, 2021
7. Bahadoer RR, Dijkstra EA, van Etten B, et al: Short-course radiotherapy followed by chemotherapy before total mesorectal excision (TME) versus preoperative chemoradiotherapy, TME, and optional adjuvant chemotherapy in locally advanced rectal cancer (RAPIDO): A randomised, open-label, phase 3 trial. Lancet Oncol 22:29-42, 2021 [Erratum: Lancet Oncol 22:e42, 2021]
8. Horvat N, Carlos Tavares Rocha C, Clemente Oliveira B, et al: MRI of rectal cancer: Tumor staging, imaging techniques, and management. Radiographics 39:367-387, 2019
9. Lambregts DMJ, Boellaard TN, Beets-Tan RGH: Response evaluation after neoadjuvant treatment for rectal cancer using modern MR imaging: A pictorial review. Insights Imaging 10:15, 2019
10. Patel UB, Taylor F, Blomqvist L, et al: Magnetic resonance imaging-detected tumor response for locally advanced rectal cancer predicts survival outcomes: MERCURY experience. J Clin Oncol 29: 3753-3760, 2011
11. Fayaz MS, Demian GA, Fathallah WM, et al: Significance of magnetic resonance imaging–assessed tumor response for locally advanced rectal cancer treated with preoperative long-course chemoradiation. J Glob Oncol 2:216-221, 2016
12. Jang J, Lee J, Park S, et al: Magnetic resonance tumour regression grade and pathological correlates in patients with rectal cancer. Br J Surg 105:1671-1679, 2018
13. Sclafani F, Brown G, Cunningham D, et al: Comparison between MRI and pathology in the assessment of tumour regression grade in rectal cancer. Br J Cancer 117:1478-1485, 2017
14. van den Broek JJ, van der Wolf FS, Lahaye MJ, et al: Accuracy of MRI in restaging locally advanced rectal cancer after preoperative chemoradiation. Dis Colon Rectum 60:274-283, 2017
15. Aker M, Boone D, Chandramohan A, et al: Diagnostic accuracy of MRI in assessing tumor regression and identifying complete response in patients with locally advanced rectal cancer after neoadjuvant treatment. Abdom Radiol (NY) 43:3213-3219, 2018
16. Lee MA, Cho SH, Seo AN, et al: Modified 3-point MRI-based tumor regression grade incorporating DWI for locally advanced rectal cancer. Am J Roentgenol 209:1247-1255, 2017
17. Jang JK, Choi SH, Park SH, et al: MR tumor regression grade for pathological complete response in rectal cancer post neoadjuvant chemoradiotherapy: A systematic review and meta-analysis for accuracy. Eur Radiol 30:2312-2323, 2020
18. George TJ Jr., Allegra CJ, Yothers G: Neoadjuvant Rectal (NAR) Score: A new surrogate endpoint in rectal cancer clinical trials. Curr Colorectal Cancer Rep 11:275-280, 2015
19. Glynne-Jones R, Glynne-Jones S: The concept and use of the neoadjuvant rectal score as a composite endpoint in rectal cancer. Lancet Oncol 22:e314-e326, 2021
20. George TJ, Yothers G, Hong TS, et al: NRG-GI002: A phase II clinical trial platform using total neoadjuvant therapy (TNT) in locally advanced rectal cancer (LARC)—First experimental arm (EA) initial results. J Clin Oncol 37:3505, 2019 (abstr TPS721)
21. Rahma OE, Yothers G, Hong TS, et al: Use of total neoadjuvant therapy for locally advanced rectal cancer: Initial results from the pembrolizumab arm of a phase 2 randomized clinical trial. JAMA Oncol 7:1225-1230, 2021
22. Chandramohan A, Siddiqi UM, Mittal R, et al: Diffusion weighted imaging improves diagnostic ability of MRI for determining complete response to neoadjuvant therapy in locally advanced rectal cancer. Eur J Radiol Open 7:100223, 2020
23. Nelson KP, Edwards D: A measure of association for ordered categorical data in population-based studies. Stat Methods Med Res 27:812-831, 2016
24. Nelson KP, Edwards D: Measures of agreement between many raters for ordinal classifications. Stat Med 34:3116-3132, 2015
25. Rullier E, Rouanet P, Tuech J-J, et al: Organ preservation for rectal cancer (GRECCAR 2): A prospective, randomised, open-label, multicentre, phase 3 trial. Lancet 390:469-479, 2017
26. Hall WA, Smith JJ: Achieving a cure without total mesorectal excision in rectal adenocarcinoma. J Clin Oncol 41:173-180, 2023
27. Chen HY, Feng LL, Li M, et al: College of American Pathologists Tumor Regression Grading System for long-term outcome in patients with locally advanced rectal cancer. Oncologist 26:e780-e793, 2021

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Prospective Correlation of Magnetic Resonance Tumor Regression Grade With Pathologic Outcomes in Total Neoadjuvant Therapy for Rectal Adenocarcinoma

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to www.asco.org/rwc or ascopubs.org/jco/authors/author-center.

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians (Open Payments).

William A. Hall
Consulting or Advisory Role: Aktis Oncology
Research Funding: Elekta (Inst)
Travel, Accommodations, Expenses: Elekta (Inst)

Marc J. Gollub
Stock and Other Ownership Interests: Pfizer

Greg Yothers
Employment: Mountainview Pediatrics
Consulting or Advisory Role: Orbus Therapeutics

Jennifer A. Dorth
Travel, Accommodations, Expenses: Varian Medical Systems

Osama E. Rahma
Employment: Outcomes4me, AstraZeneca/MedImmune
Leadership: Outcomes4me, AstraZeneca/MedImmune
Stock and Other Ownership Interests: Outcomes4Me, AstraZeneca/MedImmune
Honoraria: Merck, Clinical Care Options, MI Bioresearch, PRMA Consulting, Leerink, Alaunus Global
Consulting or Advisory Role: Celgene, Alcimed, Gfk, Merck, Five Prime Therapeutics, Putnam Associates, Defined Health, PureTech, Leerink, Genentech, Imvax, GlaxoSmithKline, Maverick Therapeutics, Bayer, Sobi
Research Funding: Amgen (Inst), Merck
Patents, Royalties, Other Intellectual Property: Pending patent (DFCI 2386.010) (Inst), PD-1/PD-L1 (Inst)
Travel, Accommodations, Expenses: Merck, Clinical Care Options, PureTech, PRMA Consulting, Genentech

Marcia M. Russell
Honoraria: Healthgrades, American College of Surgeons

Samuel A. Jacobs
Employment: Exact Sciences
Consulting or Advisory Role: Exact Sciences

Bryan A. Faller
Consulting or Advisory Role: LEK
Travel, Accommodations, Expenses: Genentech, Novartis, EB SQUIBB, Celgene, Boehringer Ingelheim, Eisai, AstraZeneca, Lilly, Amgen, Merck, Takeda
Open Payments Link: https://openpaymentsdata.cms.gov/physician/127090

Sagila George
Research Funding: Natera (Inst)
Travel, Accommodations, Expenses: Caris Life Sciences

Tareq Al baghdadi
Stock and Other Ownership Interests: Bristol Myers Squibb, Epizyme, HERON
Honoraria: Cardinal Health
Consulting or Advisory Role: Bristol Myers Squibb, Kite, a Gilead company, Lilly, AstraZeneca

Theodore S. Hong
Stock and Other Ownership Interests: PanTher Therapeutics
Consulting or Advisory Role: Merck, Synthetic Biologics, Novocure, Syndax, Boston Scientific
Research Funding: Taiho Pharmaceutical (Inst), AstraZeneca (Inst), IntraOp (Inst), Tesaro (Inst), Bristol Myers Squibb (Inst), Ipsen (Inst)

Thomas J. George
Consulting or Advisory Role: Tempus, BillionToOne, Pfizer/Array
Research Funding: Bristol Myers Squibb (Inst), Merck (Inst), AstraZeneca/MedImmune (Inst), Lilly (Inst), Bayer (Inst), Incyte (Inst), Ipsen (Inst), Seagen (Inst), Genentech (Inst), Astellas Pharma (Inst), BioMed Valley Discoveries (Inst), GlaxoSmithKline (Inst), Amgen (Inst), OncoC4 (Inst), BillionToOne (Inst), Jounce Therapeutics (Inst), Elicio Therapeutics (Inst)
Open Payments Link: https://openpaymentsdata.cms.gov/physician/321938

No other potential conflicts of interest were reported.

APPENDIX

TABLE A1. Table of Consensus MR-TRG by pCR

	pCR			
MR-TRG	Missing	0	1	Total
1	4	2	5	11
2	5	27	14	46
3	2	35	8	45
4	1	13	1	15
5	0	0	0	0
Missing	1	1	2	4
Total	13	78	30	121

Abbreviations: MR-TRG, magnetic resonance tumor regression grade; pCR, pathologic complete response.