Are your MRI contrast agents cost-effective?
Learn more about generic Gadolinium-Based Contrast Agents.
FRESENIUS KABI
caring for life

AJNR

PET/MR Imaging in Evaluating Treatment Failure of Head and Neck Malignancies: A Neck Imaging Reporting and Data System–Based Study

L.D. Patel, K. Bridgham, J. Ciriello, R. Almardawi, J. Leon, J. Hostetter, S. Yazbek and P. Raghavan

This information is current as of April 17, 2024.

AJNR Am J Neuroradiol 2022, 43 (3) 435-441
doi: https://doi.org/10.3174/ajnr.A7427
http://www.ajnr.org/content/43/3/435

PET/MR Imaging in Evaluating Treatment Failure of Head and Neck Malignancies: A Neck Imaging Reporting and Data System–Based Study

L.D. Patel, K. Bridgham, J. Ciriello, R. Almardawi, J. Leon, J. Hostetter, S. Yazbek, and P. Raghavan

ABSTRACT

BACKGROUND AND PURPOSE: PET/MR imaging is a relatively new hybrid technology that holds great promise for the evaluation of head and neck cancer. The aim of this study was to assess the performance of simultaneous PET/MR imaging versus MR imaging in the evaluation of posttreatment head and neck malignancies, as determined by its ability to predict locoregional recurrence or progression after imaging.

MATERIALS AND METHODS: The electronic medical records of patients who had posttreatment PET/MR imaging studies were reviewed, and after applying the exclusion criteria, we retrospectively included 46 studies. PET/MR imaging studies were independently reviewed by 2 neuroradiologists, who recorded scores based on the Neck Imaging Reporting and Data System (using CT/PET-CT criteria) for the diagnostic MR imaging sequences alone and the combined PET/MR imaging. Treatment failure was determined with either biopsy pathology or initiation of new treatment. Statistical analyses including univariate association, interobserver agreement, and receiver operating characteristic analysis were performed.

RESULTS: There was substantial interreader agreement among PET/MR imaging scores (κ = 0.634; 95% CI, 0.605–0.663). PET/MR imaging scores showed a strong association with treatment failure by univariate association analysis, with $P < .001$ for the primary site, neck lymph nodes, and combined sites. Receiver operating characteristic curves of PET/MR imaging scores versus treatment failure indicated statistically significant diagnostic accuracy (area under curve range, 0.864–0.987; $P < .001$).

CONCLUSIONS: Simultaneous PET/MR imaging has excellent discriminatory performance for treatment outcomes of head and neck malignancy when the Neck Imaging Reporting and Data System is applied. PET/MR imaging could play an important role in surveillance imaging for head and neck cancer.

ABBREVIATIONS: AUC = area under curve; NI-RADS = Neck Imaging Reporting and Data System; RECIST = Response Evaluation Criteria in Solid Tumors; ROC = receiver operating characteristic; SCC = squamous cell carcinoma

The importance of combining anatomic and functional imaging in the diagnosis, staging, treatment planning, and response assessment of head and neck oncology is increasingly recognized.[1,2] Hybrid PET/MR imaging is a relatively new technology that holds great promise in this regard by combining the functional evaluation of the radiotracer distribution of PET with the soft-tissue resolution of a full diagnostic neck MR imaging. MR imaging is the preferred technique for the evaluation of certain characteristics of nasopharyngeal, sinonasal, and skull base malignancies, particularly perineural spread.[3] For patients requiring surveillance of these tumors, simultaneous acquisition of a diagnostic neck MR imaging, for optimal local evaluation, and PET, for regional and distant metastasis, offers an efficient solution compared with separate PET/CT and MR imaging.

Hybrid PET/MR imaging has broad oncologic applications throughout the body, and recent studies have shown that it is often comparable with and sometimes more impactful than PET/CT.[4-6] PET/MR imaging may impact clinical management throughout diagnosis and treatment. Most comparative studies have evaluated PET/MR imaging for the initial staging of head and neck cancer in heterogeneous cohorts of predominantly squamous cell carcinoma, and most reported similar accuracy in local, nodal, and distant staging between PET/MR imaging and PET/CT.[7] A few studies reported that PET/MR imaging is advantageous for local tumor staging, specifically mentioning the benefit of T2, DWI, and contrast-enhanced sequences.[8-10] Posttreatment imaging with PET/

Received May 24, 2021; accepted after revision December 19.

From the Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine (Ringgold ID 21668), Baltimore, Maryland.

Please address correspondence Prashant Raghavan, MBBS, University of Maryland School of Medicine, 655 W Baltimore St, Baltimore, MD 21201; e-mail: praghavan@umm.edu

Indicates article with online supplemental data.

http://dx.doi.org/10.3174/ajnr.A7427

MR imaging can aid in assessing and predicting treatment response/failure; however, sample sizes have been small.[11]

PET/CT, however, remains the standard imaging technique for the evaluation of head and neck malignancies.[12] In particular, PET/CT and contrast-enhanced CT are the basis of the Neck Imaging Reporting and Data System (NI-RADS), a reporting lexicon, risk classification, and management-recommendation system for head and neck cancer surveillance imaging.[13] Standardized reporting that is data-driven and outcomes-based may increase the value that radiologists can provide in the management of patients.[14] NI-RADS categories of 1–4 at the primary tumor site and neck are based on imaging suspicion for recurrence: no evidence of recurrence, low suspicion for recurrence, high suspicion for recurrence, or known/definite recurrence.[13] Since the inception of NI-RADS in 2017, several studies have validated its use for reporting of PET/CT studies.[15-17] Most of these studies have specifically applied NI-RADS to the surveillance of squamous cell carcinoma (SCC), which accounts for most head and neck cancers.

The aim of this study was to assess the performance of simultaneous PET and diagnostic contrast-enhanced neck MR imaging in the evaluation of posttreatment head and neck malignancies, as determined by clinical outcome data after imaging. Given that NI-RADS is standardized and reproducible, we chose to apply NI-RADS to a cohort of PET/MR imaging posttreatment studies to predict residual or recurrent tumor. We perform initial surveillance imaging 8–12 weeks after treatment. Because definitive PET/MR imaging surveillance algorithms have not been established, follow-up imaging may include subsequent PET/MR imaging in 3–6 months or MR imaging alone, depending on the FDG avidity of the tumor in question or clinical decisions made at a multidisciplinary conference. Oncologists and surgeons at our institution refer both SCC and non-SCC tumors for PET/MR imaging, depending on primary tumor location. To assess PET/MR imaging performance among the entire cohort, we chose to apply NI-RADS regardless of tumor pathology, noting that most of the current literature has focused on SCC.

MATERIALS AND METHODS

This retrospective study received institutional review board approval, and the need for patient consent was waived. A data base of clinical PET/MR imaging studies performed at our institution since 2019 was searched for head and neck examinations. This search yielded 62 studies performed between May 2019 and November 2020. Review of the electronic medical records for these studies was performed to record patient age, sex, date of initial diagnosis, initial tumor staging, tumor pathology, lesion location, initial treatment, indication for PET/MR imaging, length of clinical follow-up, subsequent treatment, as well as biopsy and surgical pathology results.

Inclusion and Exclusion Criteria

Inclusion criteria were documented head and neck malignancy and an indication for PET/MR imaging for posttreatment surveillance. Criteria for treatment failure included biopsy or surgical pathology proof of residual or recurrent tumor or the initiation of new treatment (including locoregional radiation, systemic chemotherapy or palliative care). At our institution, a multidisciplinary decision to initiate new treatment without interval biopsy or surgical pathology is typically based on a combination of clinical and imaging factors suggesting disease progression, such as the Response Evaluation Criteria in Solid Tumors (RECIST 1.1; https://recist.eortc.org/) imaging criteria.[18] These criteria were used to separately designate treatment failure at the primary site and neck nodes in cases without pathology data. Lack of residual or recurrent disease was assessed by either at least 6-month disease-free clinical follow-up; at least 3-month follow-up imaging without residual tumor or recurrence; or biopsy of a suspected imaging abnormality with pathology results negative for tumor.

Exclusion criteria included insufficient outcome data to determine failure or lack of recurrence; clinically evident or pathology-proved recurrence before PET/MR imaging (NI-RADS 4); or an incomplete PET/MR imaging acquisition. Of the 50 studies initially evaluated, 4 studies were excluded according to these criteria, and 46 studies were subsequently interpreted for NI-RADS scoring.

Image Interpretation and NI-RADS Scoring

Structured reporting with NI-RADS is not the standard reporting practice at our institution. Therefore, each of the 46 studies was retrospectively interpreted and scored by 2 board-certified neuroradiologists in independent reading sessions. When scoring each study, readers first assigned a suspicion score of 1–3 based on MR images alone, followed by a NI-RADS score of 1–3 based on the entire PET/MR imaging study. The criteria used for assigning PET/MR imaging NI-RADS scores were similar to those established by the American College of Radiology for contrast-enhanced CT and PET/CT, with a focus on evaluating both FDG uptake on PET and tissue enhancement on postcontrast T1-weighted MR images. While the MR imaging acquisition included DWI and T2-weighted sequences, no specific evaluation criteria for these sequences were prescribed to the readers. Perineural spread of tumor was evaluated as a primary site finding. NI-RADS 4 "definite radiologic progression" can be subjective and difficult to define with specific criteria, especially compared with some NI-RADS 3 findings. One recommended approach to discerning NI-RADS 3-versus-4 findings is to confer with referring clinicians.[13] Therefore, in this retrospective design, readers did not assign NI-RADS 4. Separate scores were assigned to the primary tumor site and the neck; therefore, each reader assigned 4 scores to each study. Studies were reviewed by using a clinical PACS on diagnostic workstations (Carestream Health; Philips Healthcare). Consensus PET/MR imaging–based NI-RADS scores for each of the studies were assigned during a second reading session.

Image-Acquisition Methods

Simultaneous PET and MR images were acquired with an integrated PET/MR imaging system (Biograph mMR; Siemens). Imaging began 60 minutes after intravenous injection of 0.015 mCi/kg of [^{18}F] FDG. All patients had been fasting for at least 6 hours before the [^{18}F] FDG injection; fasting blood glucose levels were monitored. No patients had contraindications for MR imaging. A simultaneous whole-body PET/MR imaging acquisition was

performed, scanning across 4 bed positions from the neck through the distal thighs. Subsequently, simultaneous PET/MR imaging in the head and neck region was performed using a 20-channel head and neck receiver coil. A 2-point Dixon MR imaging sequence was acquired for attenuation correction. Additional MR images included a precontrast T1-weighted Dixon volumetric interpolated brain examination (generating images with and without fat saturation); axial and coronal T2-weighted Dixon turbo spin-echo (generating images with and without fat saturation); 3D T1-weighted spoiled gradient recalled-echo after the administration of intravenous contrast (0.1 mL/kg, gadobutrol, Gadavist; Bayer Schering Pharma); and axial DWI (b = 50 and b = 800 s/mm^2). These MR images are similar to our standard neck MR imaging protocol; however, to shorten the scan duration, we use the Dixon method to generate images with and without fat-saturation from a single acquisition; and postcontrast T1-weighted GRE sequences are a volumetric acquisition for multiplanar reformatting. The total duration of PET/MR imaging examination ranged from 45–55 minutes.

Statistical Analysis

Agreement between the NI-RADS scores assigned by the 2 readers was measured by the Cohen κ coefficient. Agreement was measured for each of the 4 scores that the 2 readers assigned to the 46 studies: primary site scored on MR imaging alone, primary site scored on PET/MR imaging, neck scored on MR imaging alone, and neck scored on PET/MR imaging. κ values were interpreted according to the commonly cited scale developed by Landis and Koch:[19] 0.01–0.20 = slight agreement, 0.21–0.40 = fair agreement, 0.41–0.60 = moderate agreement, 0.61–0.80 = substantial agreement, and 0.81–1.0 = (almost) perfect agreement.

Of the 46 PET/MR imaging studies scored for NI-RADS categories, 43 were included in statistical analyses of test performance. Three studies were subsequent PET/MR imaging studies with scores of 1 for both the primary site and neck in patients who also had an earlier study also scored as 1 at both sites. Because the treatment outcome would be the same for both studies, the subsequent study was excluded and only the initial study was incorporated into the analyses. Six patients had a second posttreatment PET/MR imaging after an operation to resect recurrent disease identified on the first posttreatment PET/MR imaging. Because the second study in these patients served as a new baseline with a different potential treatment outcome, both studies were included. Univariate association between the consensus NI-RADS category score and treatment failure was estimated using the χ^2 test and the Fisher exact test. Separate analyses were performed for the primary site, neck, and a combination of the two. To measure the performance of the consensus PET/MR imaging score to classify failure, we performed receiver operating characteristic (ROC) analysis, and the area under the curve (AUC) was calculated with 95% confidence intervals. ROC analyses and AUC were performed separately and calculated for the primary site, neck, and combination.

Because the cohort included both SCC and non-SCC tumors, a subgroup analysis of PET/MR imaging performance comparing consensus NI-RADS scores and treatment failures was performed between the SCC and non-SCC groups, with ROC analysis and the AUC. Subgroup analysis comparing the performance of the 2 readers' PET/MR imaging scores versus MR imaging–only scores was performed for the primary site and neck with ROC analysis and AUC. The statistical significance level was set at $P < .05$, and all analyses were performed using SPSS Statistics (IBM).

RESULTS

Forty-six posttreatment PET/MR imaging studies across 37 patients were evaluated. Patient demographics, tumor pathologies, and distribution of initial staging are presented in the Online Supplemental Data. The most common tumor pathology was SCC (n = 17, 45.9%). The most common lesion sites were sinonasal (n = 13, 35.1%) and nasopharyngeal (n = 6, 16.2%). At the primary site, 35.4% of lesions were T1, and 27.0% were locally advanced; 29.7% of tumors (n = 11) had evidence of perineural invasion before treatment; and 54.1% of tumors had no evidence of nodal spread at diagnosis, and most had no distant metastasis (n = 33, 89.2%).

PET/MR imaging was performed at a median of 3.2 months' posttreatment (interquartile range, 2.8–10.1 months). The median length of imaging follow-up after the initial posttreatment scan was 8.9 months (interquartile range, 6.6–12.8 months). The median length of clinical follow-up after the initial posttreatment scan was 10.8 months (interquartile range, 7.5–13.5 months).

After consensus NI-RADS scoring of all studies, 58 sites (63.0%) were assigned NI-RADS 1, 18 sites (19.6%) were assigned NI-RADS 2, and 16 sites (17.4%) were assigned NI-RADS 3. Interobserver agreement between the 2 readers for NI-RADS scores (46 studies, 92 total primary and neck sites) based on simultaneous review of PET/MR imaging was substantial at both the primary site (κ = 0.634; 95% CI, 0.605–0.663) and the neck (κ = 0.642; 95% CI, 0.611–0.673). Score agreement based solely on review of the MR image, without PET data, was low at both the primary site and neck (κ = 0.107; 95% CI, 0.0839–0.13; and κ = 0.171; 95% CI, 0.146–0.196; respectively).

The incidence of treatment failure for each NI-RADS category is listed in Table 1. Treatment failure occurred at a median of 0.9 months from the surveillance PET/MR imaging date (interquartile range, 0.5–2.5). The total incidence of failure in the cohort was 26.7% (n = 23) of 86 primary and neck sites. As detailed in the Materials and Methods section, 6 of the 92 total scored sites were excluded because they were from the 3 follow-up studies with consecutive NI-RADS scores of 1 in patients without disease recurrence. The total incidence of failure at the

Table 1: Univariate association between PET/MR imaging NI-RADS and treatment failure

PET/MR Imaging NI-RADS Score	No. Sites	Failure	*P* Value
Primary site	43	18 (41.9%)	<.001
1	19	2 (10.5%)	
2	12	5 (41.7%)	
3	12	11 (91.7%)	
Neck lymph nodes	43	5 (11.6%)	<.001
1	33	0 (0.0%)	
2	6	1 (16.7%)	
3	4	4 (100%)	
Combined sites	86	23 (26.7%)	<.001
1	52	2 (3.8%)	
2	18	6 (33.3%)	
3	16	15 (93.8%)	

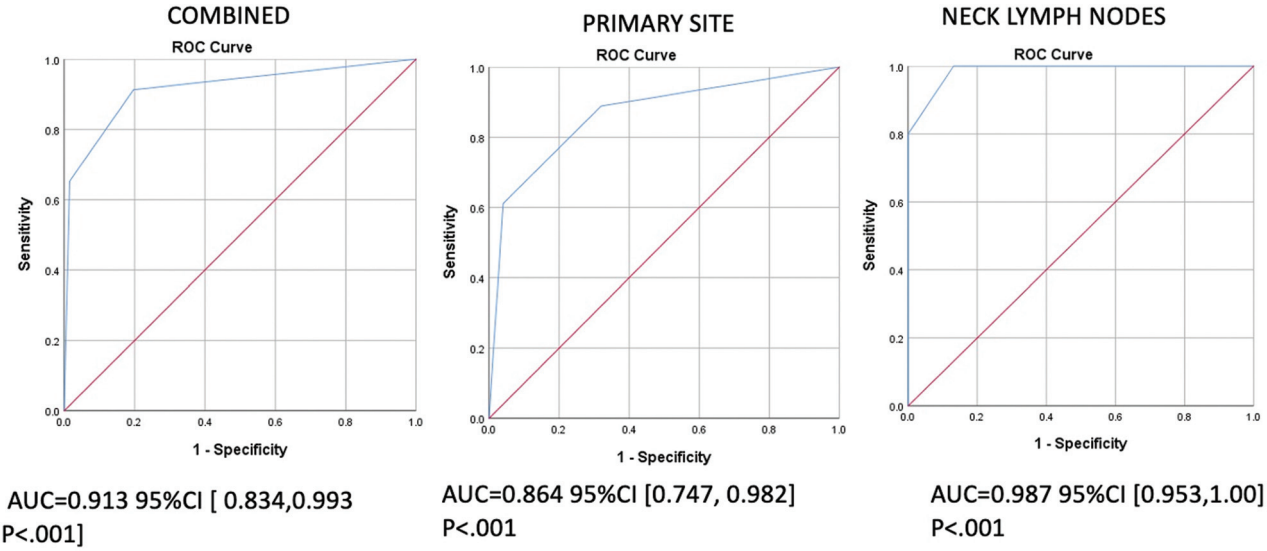

FIG 1. ROC curves for the performance of PET/MR imaging–based NI-RADS for discriminating treatment failure or no treatment failure. *Blue curves* reflect the performance of PET/MR imaging for the primary site, neck lymph nodes, and all sites combined. Inflection points in the curves are discrimination points between NI-RADS 1 and 2 and 2 and 3. The *red diagonal curve* is a reference. The AUC is reported for each curve, with values >0.500 signifying good performance of the test to discriminate treatment failure.

Table 2: Receiver operating characteristic AUC values for PET/MR imaging NI-RADS performance

ROC Curve	AUC	95% CI	*P* Value
SCC vs non-SCC subgroups			
NI-RADS 1° site vs 1° failure (SCC)	0.867	0.688–1.000	.007
NI-RADS 1° site vs 1° failure (non-SCC)	0.856	0.685–1.000	.004
NI-RADS LN vs LN failure (SCC)	1.000	1.000–1.000	.024
NI-RADS LN vs LN failure (non-SCC)	0.968	0.889–1.000	.010
PET/MR imaging NI-RADS by reader			
Reader A 1° site vs 1° failure	0.853	0.729–0.977	<.001
Reader B 1° site vs 1° failure	0.866	0.747–0.985	<.001
Reader A LN vs LN failure	0.951	0.890–1.000	<.001
Reader B LN vs LN failure	0.963	0.910–1.000	<.001
MRI only NI-RADS by reader			
Reader A 1° site vs 1° failure	0.821	0.673–0.970	<.001
Reader B 1° site vs 1° failure	0.703	0.540–0.867	.021
Reader A LN vs LN failure	0.768	0.533–1.000	.052
Reader B LN vs LN failure	0.900	0.689–1.000	.004

Note:—1° indicates the primary site; LN, lymph node.

primary site was 41.9% ($n = 18$, total 43 sites) and 11.6% at the neck ($n = 5$, total 43 sites). Eight of the primary site treatment failures had evidence of perineural spread of tumor. NI-RADS scores showed a strong association with treatment failure by univariate association analysis, with $P < .001$ for the primary site, neck lymph nodes, and combined sites. Of the 23 primary and neck sites with recurrent or residual disease, 15 were pathology-proved (65.2%). Cases without pathology were considered treatment failure due to the imaging and clinical features that led to the initiation of chemoradiation ($n = 6$, 26.1%) or palliation ($n = 2$, 8.7%). Features included progressive ulceration or lymphadenopathy on follow-up examination, skull base perineural enhancement not amenable to surgery, and progressive disease on follow-up imaging per the RECIST criteria. One case in the cohort had evidence of distant metastasis identified on PET/MR imaging.

PET/MR imaging–based NI-RADS scores showed a statistically significant performance in discriminating treatment failure. ROC curves modeling consensus NI-RADS scores against failure at the primary site, neck lymph nodes, and combined sites are presented in Fig 1. The 45° diagonal line (AUC = 0.50) corresponds to random chance; and as diagnostic accuracy of a test improves, the AUC approaches 1.0.[20] Each curve in Fig 1 has an AUC significantly greater than 0.50 ($P < .001$), indicating good performance in discriminating treatment failure versus no failure.

To compare the performance of PET/MR imaging–based NI-RADS scores in predicting treatment failure of SCC versus other malignancies, we performed a subgroup analysis. ROC curves for score performance at the primary site and neck were generated for SCC studies ($n = 19$) and non-SCC studies ($n = 24$). Performance was good in both subgroups (statistically greater AUC against reference 0.50), and there was no notable difference in the AUC for either the primary site or neck between the SCC and non-SCC groups (Table 2).

The performance of each reader's PET/MR imaging and MR imaging–only scores to discriminate failure was analyzed and compared. AUC measurements for each generated ROC curve are listed in Table 2. Each reader's MR imaging–only NI-RAD scores showed good performance, with statistically significant AUC values greater than the reference diagonal in 3 of the 4 curves and a near-significant AUC for reader A's neck scores ($P = .052$). While the AUC values of each reader's PET/MR imaging scores were higher than those of the MR imaging–only scores, these values did not reach statistical significance.

DISCUSSION

PET/MR imaging–based NI-RADS scoring demonstrated a significant association between risk categories 1, 2, and 3 and the

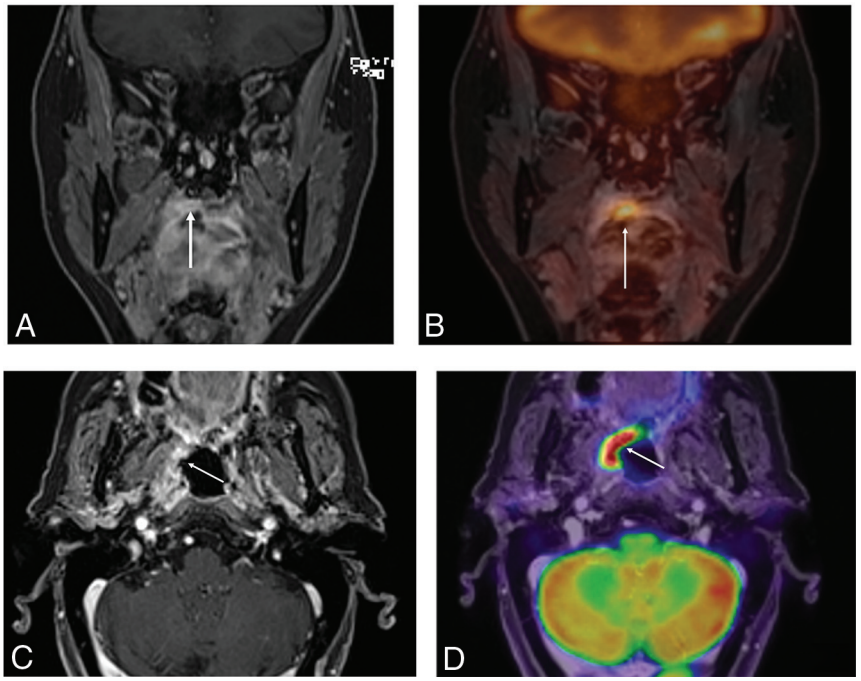

FIG 2. Conspicuity of disease recurrence with PET/MR imaging. Coronal T1WI post gadolinium (*A*) shows subtle enhancement along the right soft palate (*arrow*) in a patient with history of treated SCC. Fused coronal PET/MR imaging (*B*) demonstrates avid FDG uptake at the area of enhancement (*arrow*), which was recurrent SCC on biopsy and subsequently treated with wide local excision. Postresection axial T1WI postgadolinium sequence (*C*) shows ill-defined enhancement in the right oropharynx (*arrow*) with corresponding avid FDG uptake (*D*), which was biopsy-proved as SCC.

incidence of treatment failure in patients treated for head and neck malignancy. Higher risk scores were associated with a greater incidence of residual or recurrent disease when evaluating the primary site, neck, and all combined sites. Only 3.8% of NI-RADS 1 sites failed, compared with 33.3% and 93.8% of NI-RADS 2 and 3 sites, respectively. Furthermore, NI-RADS categories demonstrated significant performance in discriminating failure by ROC analysis, both at the primary site and neck lymph nodes. These findings suggest that PET/MR imaging–based NI-RADS scores can effectively risk-stratify patients undergoing surveillance for head and neck cancer. A NI-RADS score of 1 carries a strong negative predictive value for recurrence, whereas a score of 3 signifies high suspicion and, in this cohort, a high positive predictive value for failure. While much of the literature validating NI-RADS has focused on its application for PET/CT, this study shows a similar strong performance with PET/MR imaging.

PET/MR imaging combines the high intrinsic soft-tissue contrast of MR imaging with the metabolic information of PET imaging. The value of MR imaging in assessing features that determine tumor resectability and radiation planning, including perineural tumor spread and invasion of the orbits, skull base, vasculature, or the intracranial compartment, is well-known. These features make PET/MR imaging an attractive imaging technique to evaluate the posttreatment neck. In our study, the AUC of the ROC curves for individual readers' PET/MR imaging–based NI-RADS scores trended higher than those of the readers' MR imaging–only scores (Table 2). This finding may suggest improved performance of combined PET/MR imaging for discriminating treatment failure compared with MR imaging alone (Fig 2). This result is in keeping with other studies in which PET/MR imaging has been shown to have higher performance in diagnosing head and neck malignancy compared with PET or MR imaging alone, though NI-RADS was not applied.[21]

MR imaging is preferred to CT for the evaluation of nasopharyngeal, sinonasal, and skull base tumors, particularly those at risk for perineural spread.[3] These primary tumor locations comprised most of this PET/MR imaging cohort. Because these patients may otherwise have undergone separate MR imaging and PET/CT surveillance, PET/MR imaging likely offers a more efficient solution for patients and referring clinicians. This study cohort also highlights the potential use of PET/MR imaging and NI-RADS for surveillance of non-SCC tumors. Although SCC was the most common tumor within the cohort, most tumors were not SCC, including adenoid cystic, adenocarcinoma, and salivary gland malignancies. ROC analysis showed that PET/MR imaging–based NI-RADS performed equally well in discriminating treatment failure between the SCC and non-SCC groups. Although the existing NI-RADS literature has focused on SCC, these results suggest that NI-RADS, when applied to PET/MR imaging, could also be confidently used to evaluate other tumors. Readers should be aware of the pretreatment FDG avidity of non-SCC tumors because many have variable uptake that impacts the predictive value of PET. Adenoid cystic carcinoma, for example, has a propensity for perineural spread but can have varying FDG uptake;[22] this tumor behavior is well-suited for evaluation with PET/MR imaging. Figure 3 shows a case from this PET/MR imaging cohort of recurrent maxillary sinus adenoid cystic carcinoma with perineural spread within the orbit illustrating this feature.

Standardized reporting methods assist radiologists in providing additional value by establishing a reporting lexicon, providing linked management recommendations, and reducing interobserver variability. In this study, there was substantial interobserver agreement for NI-RADS scoring of PET/MR imaging, but only slight agreement between the readers' MR imaging–only scores. NI-RADS interpretation algorithms available from the American College of Radiology are robust in detailing how anatomic and PET data should be evaluated and reconciled.[13] These guidelines may partially account for the improved agreement between readers when scoring PET/MR imaging versus MR imaging alone. Improved reader agreement with PET/MR imaging could also indicate that interpreting MR imaging abnormalities with simultaneous, spatially coregistered PET metabolic information improves accuracy. At consensus scoring, instances of PET/MR imaging

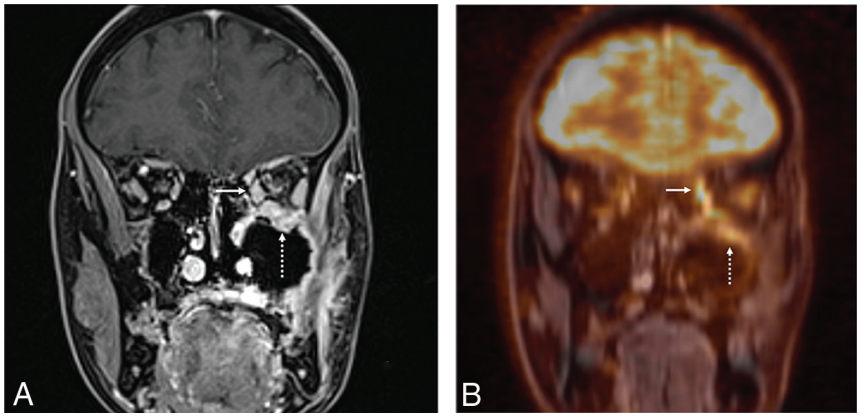

FIG 3. PET/MR imaging to evaluate perineural spread of adenoid cystic carcinoma. Coronal T1WI postgadolinium (*A*) and coronal fused PET/MR imaging (*B*) of a patient with a history of left maxillary sinus adenoid cystic carcinoma status post resection and radiation. On this initial posttreatment PET/MR imaging, the enhancing tissue along the thickened extraocular muscles (*arrow*) and infraorbital foramen (*dashed arrow*) of the left orbit corresponds to areas of increased FDG uptake and increased metabolic activity. The site was scored as NI-RADS 3 at consensus. The patient underwent orbital exenteration, with surgical pathology positive for perineural spread of tumor.

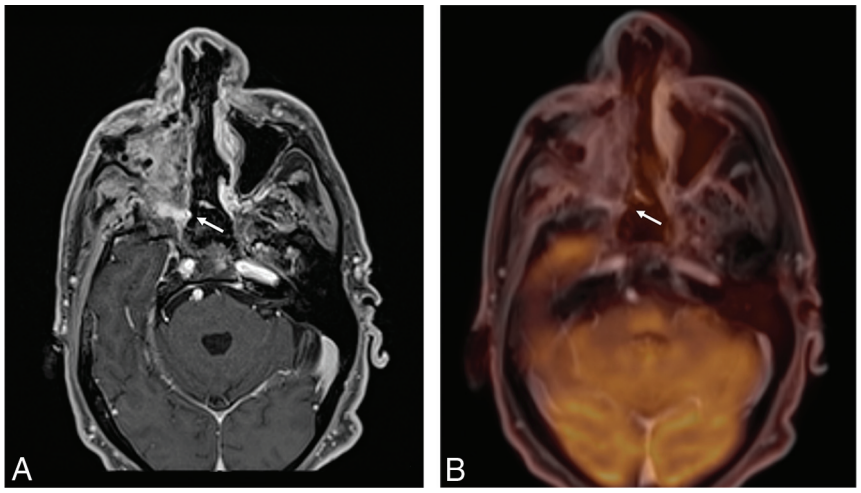

FIG 4. Interreader variability in the interpretation of PET/MR imaging–based NI-RADS. Axial T1WI postgadolinium sequence (*A*) and axial fused PET/MR imaging (*B*) in a patient with a history of maxillary sinus SCC status post subtotal maxillectomy and radiation. On the initial posttreatment PET/MR imaging, enhancing tissue in the pterygopalatine fossa (*arrow*) does not have corresponding FDG uptake. This primary site finding was scored a 2 by 1 reader (for discordant PET and MR imaging findings) and as a 1 by the other reader (for expected treatment related change). The patient has no evidence of recurrence on follow-up imaging nor clinical evidence of recurrence. In these scenarios, although agreement may not be perfect, the linked NI-RADS management decision of "short interval follow-up" allows these patients to be correctly risk-stratified.

score disagreement were mostly between consecutive NI-RADS categories, either 1 and 2 or 2 and 3. Figure 4 illustrates a sample case.

No specific MR imaging interpretation criteria were prescribed to the readers a priori, which may account for the lower agreement among MR imaging–only scores. Recent studies of NI-RADS scores applied to contrast-enhanced MR imaging report interreader agreement varying from low to substantial.[23,24] NI-RADS templates were not specifically designed for MR imaging, and recent work suggests that MR imaging features such as diffusion restriction and T2 signal be added to NI-RADS criteria,[25] possibly improving interreader agreement. Establishing MR imaging criteria may allow abbreviated MR imaging protocols and shorter image-acquisition times; this possibility could be significant, given the drawback of the time required for PET/MR imaging. In our institution, simultaneous PET/MR imaging examination time has been condensed to 45–55 minutes and includes a combination of 3D and postcontrast sequences. Recent literature has argued that gadolinium-enhanced MR images are not needed for accurate characterization of lesions because metabolic data from PET may offer similar information.[26] Thus, PET/MR imaging times could potentially be shortened even further.

We acknowledge limitations to this study. Primarily, this was not a comparative study between PET/MR imaging and PET/CT. Given differences in availability, cost, and awareness of the modalities, a comparative study is currently difficult to design. While our PET/MR imaging sample size was comparably low relative to studies of PET/CT–based NI-RADS, it is similar to that in other studies evaluating PET/MR imaging. As clinical use of PET/MR imaging increases,[7] it may be possible to power a study designed to compare PET/CT and PET/MR imaging. Another limitation is that PET/MR imaging NI-RADS scoring was performed in a retrospective manner because NI-RADS is not a standard reporting practice at our institution. However, a potential outcome of this method is that the interobserver agreement we observed may be more representative of general practice, in which use of NI-RADS likely is still expanding. Furthermore, NI-RADS 4 scores were not retrospectively assigned, because definitive radiologic disease progression can be subjective and difficult to discern from NI-RADS 3 findings such as a discrete mass with intense FDG uptake. This methodology may have introduced sampling bias within the NI-RADS 3 cohort and could account for the higher failure rate among the NI-RADS 3 scores in this study compared with others.

Although 8 cases in this cohort lacked pathologic proof of recurrence, the multidisciplinary decision to initiate treatment

without a biopsy was in response to imaging and clinical features. A false-positive case is difficult to exclude in this scenario because without tissue diagnosis, clinical features such as progressive ulceration could be the sequelae of radiation rather than true recurrence. A combination of imaging and clinical criteria could be defined for the NI-RADS 4 category, though this would need to be institution- and practice-specific.

CONCLUSIONS

Hybrid imaging with simultaneous PET/MR imaging may offer a more patient-friendly alternative to sequential MR imaging and PET/CT imaging, while maintaining excellent diagnostic performance and, as shown in this study, excellent discriminatory performance for treatment failure when NI-RADS is applied. PET/MR imaging could play an important role in surveillance imaging for head and neck cancer, depending on the site of the primary tumor and the particular pathology. Use of standardized reporting and management recommendations such as NI-RADS may make PET/MR imaging more appealing to oncologists and surgeons.

ACKNOWLEDGMENTS

The authors would like to acknowledge Brigitte Pocta for her assistance in manuscript editing and preparation of visuals and John Hebel for his contribution to PET/MR imaging protocols and image-acquisition methodology.

Disclosure forms provided by the authors are available with the full text and PDF of this article at www.ajnr.org.

REFERENCES

1. Hohenstein NA, Chan JW, Wu SY, et al. **Diagnosis, staging, radiation treatment response assessment, and outcome prognostication of head and neck cancers using PET imaging: a systematic review.** *PET Clin* 2020;15:65–75 CrossRef Medline
2. Özel HE. **Use of PET in head and neck cancers.** *Turk Arch Otorhinolaryngol* 2015;53:73–76 CrossRef Medline
3. Saito N, Nadgir RN, Nakahira M, et al. **Posttreatment CT and MR imaging in head and neck cancer: what the radiologist needs to know.** *Radiographics* 2012;32:1261–82; discussion 1282–84 CrossRef Medline
4. Bashir U, Mallia A, Stirling J, et al. **PET/MRI in oncological imaging: state of the art.** *Diagnostics (Basel)* 2015;5:333–57 CrossRef Medline
5. Broski SM, Goenka AH, Kemp BJ, et al. **Clinical PET/MRI: 2018 update.** *AJR Am J Roentgenol* 2018;211:295–313 CrossRef Medline
6. Ehman EC, Johnson GB, Villanueva-Meyer JE, et al. **PET/MRI: where might it replace PET/CT?** *J Magn Reson Imaging* 2017;46:1247–62 CrossRef Medline
7. Huellner MW. **PET/MR imaging in head and neck cancer: an update.** *Semin Nucl Med* 2021;51:26–38 CrossRef Medline
8. Samolyk-Kogaczewska N, Sierko E, Dziemianczyk-Pakiela D, et al. **Usefulness of hybrid PET/MRI in clinical evaluation of head and neck cancer patients.** *Cancers (Basel)* 2020;12:511 CrossRef Medline
9. Huang SH, Chien CY, Lin WC, et al. **A comparative study of fused FDG PET/MRI, PET/CT, MRI, and CT imaging for assessing surrounding tissue invasion of advanced buccal squamous cell carcinoma.** *Clin Nucl Med* 2011;36:518–25 Jul CrossRef Medline
10. Cheng Y, Bai L, Shang J, et al. **Preliminary clinical results for PET/MR imaging compared with PET/CT in patients with nasopharyngeal carcinoma.** *Oncol Rep* 2020;43:177–87 CrossRef Medline
11. Romeo V, Iorio B, Mesolella M, et al. **Simultaneous PET/MRI in assessing the response to chemo/radiotherapy in head and neck carcinoma: initial experience.** *Med Oncol* 2018;35:112 CrossRef Medline
12. Goel R, Moore W, Sumer B, et al. **Clinical practice in PET/CT for the management of head and neck squamous cell cancer.** *AJR Am J Roentgenol* 2017;209:289–303 CrossRef Medline
13. Aiken AH, Rath TJ, Anzai Y, et al. **ACR Neck Imaging Reporting and Data Systems (NI-RADS): a white paper of the ACR NI-RADS Committee.** *J Am Coll Radiol* 2018;15:1097–1108 CrossRef Medline
14. Dodd GD, Allen B, Birzniek D, et al. **Reengineering the radiology enterprise: a summary of the 2014 Intersociety Committee Summer Conference.** *J Am Coll Radiol* 2015;12:228–34 CrossRef Medline
15. Krieger DA, Hudgins PA, Nayak GK, et al. **Initial performance of NI-RADS to predict residual or recurrent head and neck squamous cell carcinoma.** *AJNR Am J Neuroradiol* 2017;38:1193–99 CrossRef Medline
16. Wangaryattawanich P, Branstetter BF, Hughes M, et al. **Negative predictive value of NI-RADS category 2 in the first posttreatment FDG-PET/CT in head and neck squamous cell carcinoma.** *AJNR Am J Neuroradiol* 2018;39:1884–88 CrossRef Medline
17. Wangaryattawanich P, Branstetter BF, Ly JD, et al. **Positive predictive value of neck imaging reporting and data system categories 3 and 4 posttreatment FDG-PET/CT in head and neck squamous cell carcinoma.** *AJNR Am J Neuroradiol* 2020;41:1070–75 CrossRef Medline
18. Eisenhauer EA, Therasse P, Bogaerts J, et al. **New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1).** *Eur J Cancer* 2009;45:228–47 CrossRef Medline
19. Landis JR, Koch GG. **The measurement of observer agreement for categorical data.** *Biometrics* 1977;33:159–74 Medline
20. Zou KH, O'Malley AJ, Mauri L. **Receiver-operating characteristic analysis for evaluating diagnostic tests and predictive models.** *Circulation* 2007;115:654–57 CrossRef Medline
21. Park J, Pak K, Yun TJ, et al. **Diagnostic accuracy and confidence of [18F] FDG PET/MRI in comparison with PET or MRI alone in head and neck cancer.** *Sci Rep* 2020;10:9490 CrossRef Medline
22. Ruhlmann V, Poeppel TD, Veit J, et al. **Diagnostic accuracy of 18 F-FDG PET/CT and MR imaging in patients with adenoid cystic carcinoma.** *BMC Cancer* 2017;17:887 CrossRef Medline
23. Elsholtz FH, Erxleben C, Bauknecht HC, et al. **Reliability of NI-RADS criteria in the interpretation of contrast-enhanced magnetic resonance imaging considering the potential role of diffusion-weighted imaging.** *Eur Radiol* 2021;31:6295–6304 CrossRef Medline
24. Abdelaziz TT, Abdel Razk AK, Ashour MM, et al. **Interreader reproducibility of the Neck Imaging Reporting and Data System (NI-RADS) lexicon for the detection of residual/recurrent disease in treated head and neck squamous cell carcinoma (HNSCC).** *Cancer Imaging* 2020;20:61 CrossRef Medline
25. Ashour MM, Darwish EA, Fahiem RM, et al. **MRI posttreatment surveillance for head and neck squamous cell carcinoma: proposed MR NI-RADS criteria.** *AJNR Am J Neuroradiol* 2021;42:1123–29 CrossRef Medline
26. Pyatigorskaya N, De Laroche R, Bera G, et al. **Are gadolinium-enhanced MR sequences needed in simultaneous 18F-FDG-PET/MRI for tumor delineation in head and neck cancer?** *AJNR Am J Neuroradiol* 2020;41:1888–96 CrossRef Medline