

Initial Performance of NI-RADS to Predict Residual or Recurrent Head and Neck Squamous Cell Carcinoma

^{ID}D.A. Krieger, ^{ID}P.A. Hudgins, ^{ID}G.K. Nayak, ^{ID}K.L. Baugnon, ^{ID}A.S. Corey, ^{ID}M.R. Patel, ^{ID}J.J. Beitler, ^{ID}N.F. Saba, ^{ID}Y. Liu, and ^{ID}A.H. Aiken

ABSTRACT

BACKGROUND AND PURPOSE: The Head and Neck Imaging Reporting and Data System (NI-RADS) surveillance template for head and neck cancer includes a numeric assessment of suspicion for recurrence (1–4) for the primary site and neck. Category 1 indicates no evidence of recurrence; category 2, low suspicion of recurrence; category 3, high suspicion of recurrence; and category 4, known recurrence. Our purpose was to evaluate the performance of the NI-RADS scoring system to predict local and regional disease recurrence or persistence.

MATERIALS AND METHODS: This study was classified as a quality-improvement project by the institutional review board. A retrospective database search yielded 500 consecutive cases interpreted using the NI-RADS template. Cases without a numeric score, non-squamous cell carcinoma primary tumors, and primary squamous cell carcinoma outside the head and neck were excluded. The electronic medical record was reviewed to determine the subsequent management, pathology results, and outcome of clinical and radiologic follow-up.

RESULTS: A total of 318 scans and 618 targets (314 primary targets and 304 nodal targets) met the inclusion criteria. Among the 618 targets, 85.4% were scored NI-RADS 1; 9.4% were scored NI-RADS 2; and 5.2% were scored NI-RADS 3. The rates of positive disease were 3.79%, 17.2%, and 59.4% for each NI-RADS category, respectively. Univariate association analysis demonstrated a strong association between the NI-RADS score and ultimate disease recurrence, with $P < .001$ for primary and regional sites.

CONCLUSIONS: The baseline performance of NI-RADS was good, demonstrating significant discrimination among the categories 1–3 for predicting disease.

ABBREVIATIONS: AUC = area under the curve; CECT = contrast-enhanced CT; H&N = head and neck; NI-RADS = Head and Neck Imaging Reporting and Data System; ROC = receiver operating characteristic

Radiologists are major stake-holders in the shift toward value-based performance. The American College of Radiology is leading the effort to re-engineer the radiology enterprise to be “patient centric, data-driven, and outcomes-based.” Standard-

ized reporting systems, dictation templates, and linked management recommendations have been identified as key contributions to value.¹ Much of this shift toward data-driven and outcomes-based reporting stems from the success of the BI-RADS system for standardizing mammography reports. Similar templates have been developed for hepatocellular carcinoma,² prostate cancer,³ and thyroid nodules.⁴

The Head and Neck Imaging Reporting and Data System (NI-RADS) was recently developed for surveillance contrast-enhanced CT (CECT) with and without positron-emission tomography in patients with treated head and neck (H&N) cancer.⁵ Both the primary tumor site and neck are assessed for recurrence and assigned a category of 1–4 based on imaging suspicion with linked management recommendations:

- Category 1: No evidence of recurrence
Imaging: Expected posttreatment change (tissue distortion, scar, radiation change)

Received November 22, 2016; accepted after revision January 22, 2017.

From the Department of Radiology, Hackensack University Medical Center (D.A.K.), Hackensack, New Jersey; Departments of Radiology and Imaging Sciences (P.A.H., A.H.A., K.L.B., A.S.C.) and Otolaryngology Head and Neck Surgery (M.R.P., J.J.B.), Emory University School of Medicine, Atlanta, Georgia; Departments of Radiation Oncology (J.J.B.), Hematology and Medical Oncology (J.J.B., N.F.S.), and Biostatistics (Y.L.), Winship Cancer Institute, Emory University School of Medicine, Atlanta, Georgia; and Department of Radiology (G.K.N.), New York University School of Medicine, New York, New York.

Paper previously presented as an oral presentation at: Annual Meeting of the American Society of Neuroradiology and the Foundation of the ASNR Symposium, May 21–26, 2016; Washington, DC.

Please address correspondence to Ashley Aiken, MD, Emory University Hospital, 1364 Clifton Rd NE, Atlanta, GA 30322; e-mail: Ashley.aiken@emoryhealthcare.org; @phudge54

<http://dx.doi.org/10.3174/ajnr.A5157>

Management: Routine surveillance (6 months typically, see “Materials and Methods”)

- Category 2: Low suspicion of recurrence
Imaging: Ill-defined abnormality with only mild enhancement and/or FDG uptake
Management: Direct inspection for mucosal findings or short (3-month) follow-up with CECT or PET for deep findings
- Category 3: High suspicion of recurrence
Imaging: Discrete, new, or enlarging lesions with intense enhancement and/or focal FDG uptake
Management: Biopsy
- Category 4: Known recurrence, pathologically proved or definite radiologic or clinical progression.

NI-RADS categories 1–4 are the same for CECT or CECT/PET, but the linked management recommendations and lexicon are slightly different (because FDG avidity is included in the latter). Furthermore, the first version of NI-RADS category 2 contains subcategories that also address lesion size and location (superficial or deep): 2a, superficial/mucosal surface; 2b, deep abnormality of <1 cm; and 2c, deep abnormality of >1 cm.⁵ These subcategories are useful to direct management, for example superficial mucosal abnormalities are amenable to direct inspection. Size criteria were not added to predict disease but rather to avoid biopsy in this indeterminate category unless immediate management depended on the biopsy.⁵ This template-driven approach reflects common language to promote collaboration between radiologists and referring providers, data-driven optimization of H&N cancer imaging, and greater direct engagement with patients.

An obstacle to improvement in value-based performance and direct patient reporting is lack of a data-driven standard surveillance imaging algorithm. PET/CECT at 12 weeks is often the first posttreatment study, though a recent study suggests that it can be performed at 8 weeks.⁶ At our institution, patients with advanced H&N cancer are scanned with CECT/PET at 12 weeks as a baseline. If the findings are negative, they undergo CECT alone 6 months later, and if these findings are negative, they undergo CECT alone 12 months later. Although studies have investigated PET/CT for surveillance,^{7–9} ordering practices among treating physicians remain variable. The 2015 National Comprehensive Cancer Network recommendations advocate imaging within 6 months for T3/4 primary tumors or N2/3 nodal disease and then additional imaging only for new signs/symptoms, smoking, or areas inaccessible to clinical inspection (the latter being arbitrary and difficult to apply).¹⁰ Yet, 79% of H&N cancer surgeons self-reported using PET/CT for asymptomatic patients.¹¹ Given this variation in practice, it is critical to have measurable categories to correlate with outcomes to develop a data-driven universal surveillance algorithm.

NI-RADS allows H&N radiologists to perform structured radiologic-pathologic correlation and to determine accuracy, prognostic value, and interobserver agreement in contrast to subjective interpretations that provide no data for retrospective analysis. Our objective was to determine the initial performance of the NI-RADS scoring system to predict tumor recurrence or persistence in patients treated for squamous cell carcinoma of the H&N undergoing imaging surveillance.

MATERIALS AND METHODS

This study was designated a Quality Improvement project by our institutional review board at the Emory University School of Medicine. An electronic medical record search from June 12, 2014 to January 28, 2015 yielded 500 consecutive neck CECT examinations interpreted with the NI-RADS template, including patients with a variety of tumor types and primary locations. The following was gathered from a review of the electronic medical record:

- 1) Age and sex
- 2) Date and site of original diagnosis
- 3) Human papillomavirus status
- 4) Initial tumor stage (Tumor, Node, Metastasis)
- 5) Treatment (ie, surgery, chemotherapy, radiation)
- 6) Date of the index scan
- 7) Type of index scan (CECT alone versus CECT with PET scan)
- 8) First posttreatment scan versus subsequent surveillance scan
- 9) Length of imaging and clinical follow-up

Inclusion Criteria

- 1) Treated primary H&N squamous cell carcinoma.
- 2) CECT and/or CECT/PET for surveillance.
- 3) NI-RADS template used for interpretation.

A total of 402 scans met the inclusion criteria.

Criteria for tumor recurrence or persistence included the following: 1) Biopsy positive for squamous cell carcinoma, 2) evidence of disease progression on subsequent imaging (per Response Evaluation Criteria In Solid Tumors criteria; <http://www.radiologytutor.com/index.php/cases/oncol/139-recist>), or 3) obvious tumor on physical examination. To confirm the lack of tumor recurrence, we assessed the following: 1) follow-up imaging at least 90 days after the index scan, 2) clinical follow-up for at least 6 months without evidence of recurrence, or 3) biopsy of an abnormality detected on the index scan with pathology results negative for tumor.

Exclusion Criteria

- 1) Insufficient outcomes data to determine positive or negative disease.
- 2) Score of “4, known recurrence” because recurrence had already been proved before the scan. However, it is possible to have a score of 4 for the primary site and 1, 2, or 3 for lymph nodes (or vice versa). Thus, an outcome could still be determined for nodes so that each scan had 2 possible sites for target abnormalities (primary and neck).
- 3) Multiple scans in the same patient if there were back-to-back scores of 1 for both primary and neck. In this case, subsequent index scans were excluded because the final outcome of “recurrence or not” for the primary or neck would be the same for these 2 data points.

These criteria yielded 287 patients, 318 scans, and 618 total targets (314 primary targets and 304 nodal targets) for which outcomes could be determined.

Surveillance Algorithm and Image Interpretation

At our institution, all patients with advanced H&N cancer (almost all patients except those with T1 N0 disease) are scanned with CECT/PET at 12-week baseline, and if the findings are negative, they undergo a CECT alone 6 months later. If these findings are negative, they undergo a CECT 12 months later. All NI-RADS surveillance scans were interpreted prospectively using the template by 1 of 4 dedicated H&N neuroradiologists (30, 15, 10, and 9

years of experience). Both the primary site and neck were assigned a NI-RADS category of 1–4. For this study, all category 2 subcategories were recorded as general category 2. For scores of 2–4, the target abnormality was described briefly in the impression after the numeric score. The NI-RADS template, created by a multidisciplinary team and implemented in 2014, has been subject to ongoing peer review through weekly tumor boards and the American College of Radiology RADPEER. Interpreting radiologists reviewed prior clinical history and endoscopic notes. Comparison with baseline imaging, including pretreatment FDG avidity when available, was made. The subjective interpretation of the PET/CECT included evaluation of disease on both fused PET and CECT. As noted in the NI-RADS template, factors incorporated into lesion assessment included the following: size, FDG avidity, morphology, and enhancement pattern. Because previous studies have established that the standard uptake value data do not improve diagnostic accuracy for disease after treatment for H&N cancer, a strict threshold for standard uptake value was not used.^{6,12,13} Instead, a subjective dichotomous analysis of intense FDG uptake was determined.

Image Acquisition

All PET/CT imaging followed standard protocol, and was performed on GE Discover 600 and 690 PET/CT scanners (GE Healthcare, Milwaukee, Wisconsin). Patients fasted for 6 hours before the scan, and serum glucose concentration was obtained immediately before FDG administration. The examination was deferred if glucose was >200 mg/dL. Combined PET/CT from the skull vertex through the midhigh was obtained 1 hour after intravenous administration of 10–14 mCi of FDG. Helical noncontrast CT from the vertex through midhigh was performed before PET for attenuation correction and anatomic localization. A CECT of the neck with the arms down was performed following PET. Our split-bolus technique used 110 mL of intravenous iopamidol (Isovue-370; Bracco, Princeton, New Jersey), with 55 mL injected first at 2.5 mL/s, a 40-second delay, then another 55 mL at the same rate, with a total scan delay of 90 seconds. We acquired axial images from the frontal sinuses through the mediastinum at 1.25-mm section thickness; pitch, 0.984:1; gantry rotation, 0.7 seconds; FOV, 25 cm; 120 kV(peak); and Smart milliampere with a noise index of 13.78. Reformatted images at 2.5-mm thickness in the axial planes and 3-mm sagittal and coronal reformations were sent to the PACS.

Statistical Methods

Univariate association between recurrence and scan score (1–3) was estimated by the χ^2 test and the nonparametric Fisher exact test. The same analysis was repeated for primary site, lymph node, and their combination separately. The overall performance of discrimination of the scan score on recurrence status (yes versus no) is measured as the area under curve (AUC) by receiver operating characteristic (ROC) analysis with 95% confidence intervals. The sensitivity and specificity at each cut-point of the scan score were presented accordingly for score 1 versus 2–3 and for scores 1–2 versus 3. Additionally, the same ROC analyses for subgroup performed for CECT alone versus CECT + PET and for the first posttreatment examination versus the subsequent surveillance

Table 1: Tumor site and initial stage (patient level)

Site/Stage	% (No.)
Primary site	
Oropharynx	43.2% (124)
Larynx	22.3% (64)
Oral cavity	25.4% (73)
Hypopharynx	4.2% (12)
Skin	2.1% (6)
Unknown	2.8% (8)
Primary stage	
Tx	7.7% (22)
Tis	0.3% (1)
T1	16.4% (47)
T2	25.4% (73)
T3	12.2% (35)
T4a	32.8% (94)
T4b	4.2% (12)
T4	0.8% (3)
Nodal stage	
Nx	4.5% (13)
N0	28.9% (83)
N1	11.8% (34)
N2a	3.1% (9)
N2b	33.8% (97)
N2c	16.4% (47)
N3	1.4% (4)
Distant stage	
M0	97.9% (281)
M1	2.1% (6)

examination were explored. The interobserver agreement was measured by κ statistics among 40 scans for primary and neck sites by 2 graders. The statistical significance level was set at $P < .05$, and analyses were conducted in SAS 9.4 (SAS Institute, Cary, North Carolina).

RESULTS

Of the 318 examinations, there were 221 CECTs alone (69.5%) and 97 CECT/PETs (30.5%). Sixty studies (16.4%) were initial baseline posttreatment examinations (performed at 12 weeks); the remainder were follow-up examinations during routine surveillance per protocol.

Median imaging follow-up after the index scan was 51 weeks; median clinical follow-up was 54 weeks. The distribution of tumor site and initial stage (when known) is outlined in Table 1. Primary tumors of the oropharynx were the largest group (43.2%), followed by tumors of the oral cavity (25.4%) and larynx (22.3%). At the primary site, almost one-third had moderately advanced (T4a) disease (32.8%). More than half had at least N2 nodal disease (54.7%). Distant metastatic disease at initial staging was rare (2.1%).

Interobserver Agreement

The interobserver agreement determined by κ statistics after review of 40 scans (80 targets) by 2 graders was very good, 0.821 (95% CI, 0.657–0.986) with $P < .001$.

Incidence of Disease Recurrence/Persistence Based on NI-RADS Score

The incidence of recurrence for each NI-RADS category is detailed in Table 2. Overall, the incidence of tumor persistence/recurrence was 7.9%, with an 8.9% (28/314) recurrence rate at the primary site and a 6.9% (21/304) regional nodal recurrence rate.

Table 2: Recurrence rates among the NI-RADS categories

NI-RADS Categories	Total	Recurrence Rate (No.)
Primary site		
NI-RADS 1	254	3.5% (9)
NI-RADS 2	38	18.4% (7)
NI-RADS 3	22	54.6% (12)
All primary site categories	314	8.9% (28)
Lymph nodes		
NI-RADS 1	274	4.0% (11)
NI-RADS 2	20	15.0% (3)
NI-RADS 3	10	70.0% (7)
All nodal categories	304	6.9% (21)
Combined primary and nodes		
NI-RADS 1	528	3.8% (20)
NI-RADS 2	58	17.2% (10)
NI-RADS 3	32	59.4% (19)
Combined, all categories	618	7.9% (49)

NI-RADS 1. Five hundred twenty-eight of 618 targets (85.4%) were scored “NI-RADS 1, no evidence of recurrence” with only 3.8% having recurrent disease during the follow-up. When considered separately, the recurrence rate for primary and nodal NI-RADS 1 scores was similar (3.5% and 4.0%, respectively).

NI-RADS 2. Fifty-eight of 618 targets (9.4%) were scored “NI-RADS 2, questionable recurrence” and had a higher overall rate of recurrence of 17.2%, with similar rates for primary and nodes separately (18.4% versus 15.0%). Of 58/618 category 2 lesions, there were 38 primary site category 2 lesions (27/38 “2a,” 7/38 “2b,” and 4/38 “2c”). Seven of 38 underwent biopsy with 5/7 positive, and 2 patients had imaging progression, for a total recurrence of 7/38 (18.4%). There were 20 neck category 2 lesions (15 “2a” and 5 “2b”) with 2/20 having pathology-proven recurrence and 1/20 with clinical disease progression, for a total recurrence rate of 3/20 (15.0%) for the neck. There was no difference in the rate of positive disease based on lesion size within the confines of this small sample size.

NI-RADS 3. Thirty-two of 618 targets (5.2%) were “NI-RADS 3, highly suspicious for recurrence” and had the highest overall recurrence rate of 59.4%, with a 54.6% recurrence at the primary site and a 70.0% rate at nodes. Of the 32/618 category 3 lesions, there were 22 primary site lesions and 10 neck lesions. Twenty-two of 32 category 3 targets had pathologic confirmation of disease presence or absence. The remaining 10/32 did not have pathologic confirmation because it would not affect management ($n = 7$) or the ultrasound or CT correlate of the suspected lesion could not be found when biopsy was attempted ($n = 3$). Eight of these 10 (80.0%) had clinical or radiologic evidence for recurrence (7 primary site lesions and 1 nodal site), defined as progression at the target site on imaging or clinically obvious tumor.

NI-RADS Performance

Univariate association analysis demonstrated a strong association between the NI-RADS score and ultimate disease persistence/recurrence, with $P < .001$ for primary site, lymph node scores, and combined scores. ROC curves for NI-RADS performance at the primary site (Fig 1), lymph nodes (Fig 2), and combined (Fig 3) were obtained and reflect an overall good performance. For the primary site ROC curve (Fig 1) (AUC = 0.787; 95% CI, 0.691–

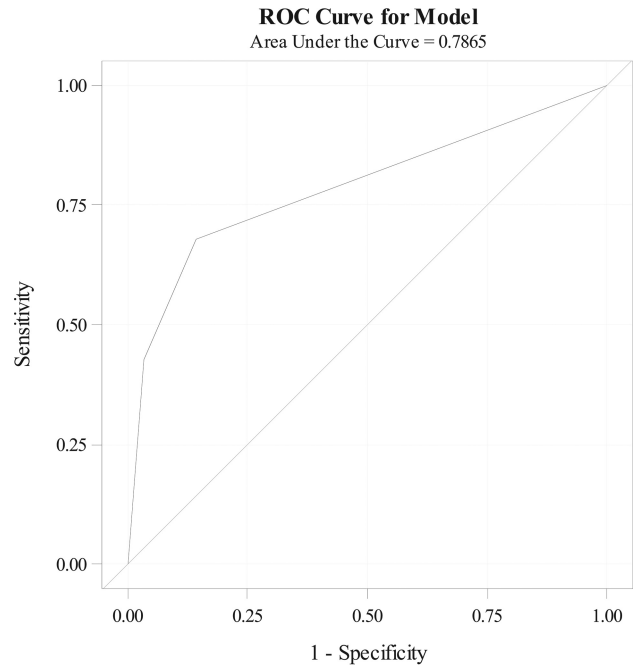


FIG 1. ROC curve for NI-RADS at the primary site with AUC = 0.786 (95% CI, 0.691–0.881).

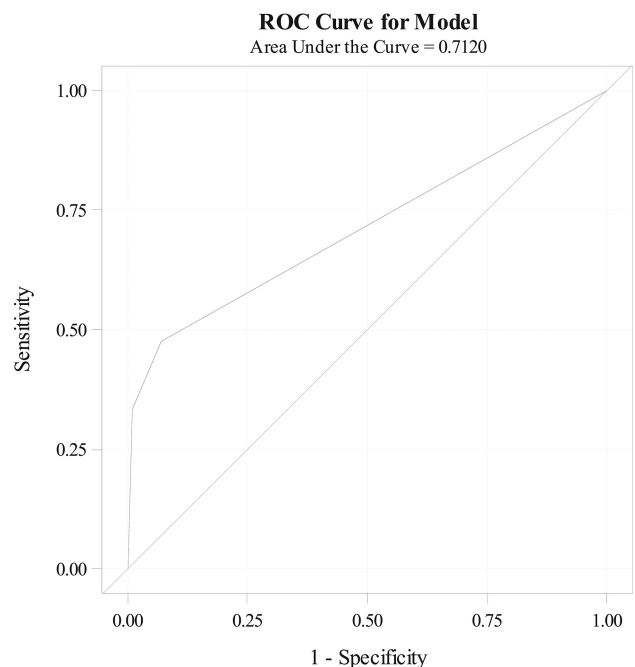


FIG 2. ROC curve for NI-RADS at the lymph nodes with AUC = 0.71 (95% CI, 0.597–0.826).

0.881), $P < .001$ indicated a good performance of the NI-RADS score to discriminate primary site recurrence versus no recurrence (an AUC value of 1 indicates a perfect discrimination, and an AUC value of 0.5 indicates no use). For lymph nodes, the AUC of 0.712 and an AUC of 0.756 for combined primary and nodal sites indicated good overall performance of this rating scale.

Subgroup Analysis of CECT Alone versus PET/CECT

A subgroup analysis was undertaken comparing the performance of CECT alone versus CECT + PET/CT (Table 3). The overall

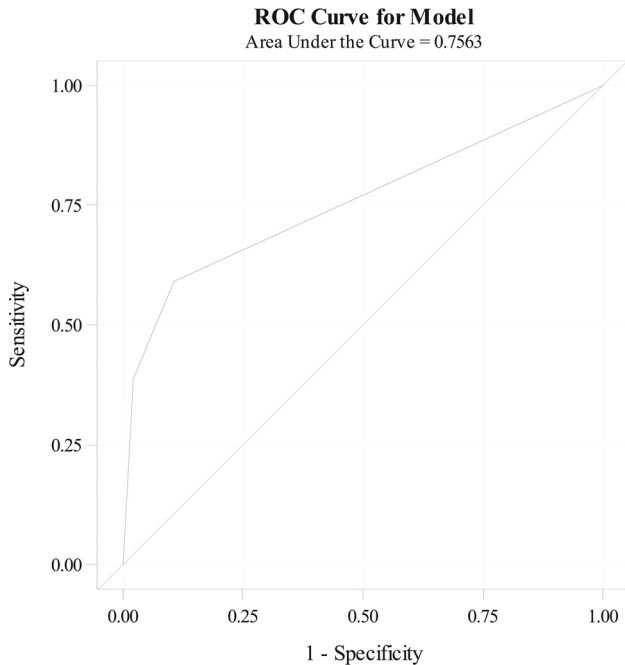


FIG 3. ROC curve for NI-RADS for primary site and lymph nodes combined, with AUC = 0.756 (95% CI, 0.682–0.8).

Table 3: CECT alone versus CECT with PET/CT

	CECT	CECT + PET/CT
Combined primary and nodes		
NI-RADS 1	3.1% (12/385)	5.6% (8/143)
NI-RADS 2	21.9% (7/32)	11.5% (3/26)
NI-RADS 3	91.7% (11/12)	40.0% (8/20)
Combined, all categories	7.0% (30/429)	10.1% (19/189)

Table 4: Initial posttreatment versus subsequent follow-up

Combined Primary and Nodes	Posttreatment	Follow-Up
NI-RADS 1	5.7% (5/88)	3.4% (15/440)
NI-RADS 2	20.0% (4/20)	15.8% (6/38)
NI-RADS 3	50.0% (4/8)	62.5% (15/24)
Combined, all categories	11.2% (13/116)	7.2% (36/502)

recurrence rate in these 2 groups was similar (7.0% versus 10.1%). Although there was no statistical difference in overall performance of NI-RADS for CECT (AUC = 0.779) versus CECT/PET (AUC = 0.709), a NI-RADS 3 on CECT alone was more likely to correctly identify recurrence (primary or nodal) compared with a NI-RADS 3 on CECT + PET (91.7% versus 40.0%).

Subgroup Analysis of Initial Posttreatment Study versus Subsequent Studies

An additional subgroup analysis compared the performance on initial posttreatment studies with performance on subsequent follow-up (Table 4). While there was no statistical difference in the overall performance of NI-RADS in initial posttreatment surveillance (AUC = 0.729) versus subsequent scans (AUC = 0.760), the recurrence rate for NI-RADS 1 was greater for the initial baseline scan group (5.7%, 5/88) compared with the subsequent follow-up examination group (3.4%, 15/440). This difference was even more pronounced when looking at the primary site alone (9.3% versus 2.4%). As expected, the incidence of positive disease was

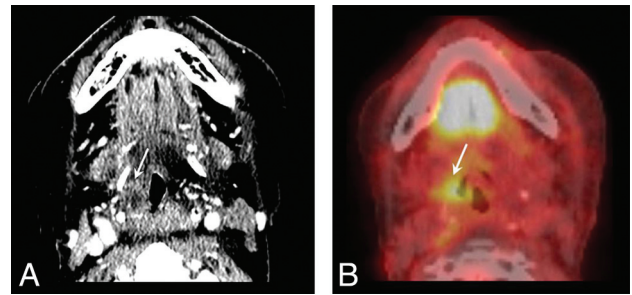


FIG 4. NI-RADS primary site category 2a: superficial mucosal abnormality. Primary T4a N2c base of tongue squamous cell carcinoma, status post chemoradiotherapy. A, CECT showed only subtle/questionable asymmetric enhancement in the right vallecula (arrow) retrospectively after review of PET. B, Fused PET image shows asymmetric uptake in the right vallecula (arrow). Direct visualization did show ulcerated mucosa, but the biopsy was negative for tumor. Clinically, this was deemed a radiation-related injury.

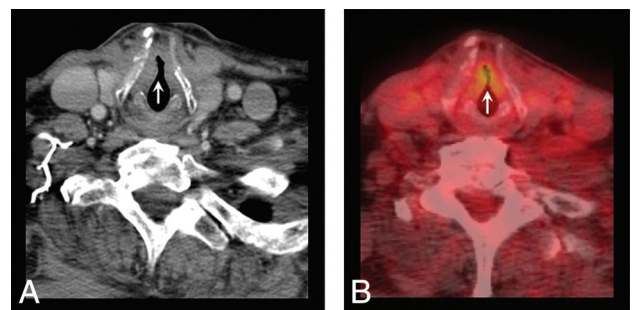


FIG 5. NI-RADS primary site category 2a: superficial mucosal abnormality. Primary T2 larynx squamous cell carcinoma status post chemoradiotherapy. A, CECT showed subtle irregularity of the anterior commissure and anterior true vocal cords bilaterally (arrow). B, Corresponding fused PET image shows focal mucosal uptake (arrow). After direct visualization revealed suspicious mucosal findings, the biopsy showed persistent disease. Although this lesion does demonstrate focal avid FDG uptake, it is in a special category of mucosal abnormality. In the published NI-RADS 1.0 by Aiken et al,⁵ these are scored as 2a because the linked management recommendation is direct visualization.

also greater in the initial posttreatment group versus surveillance studies (11.2% versus 7.2%).

DISCUSSION

The baseline performance of NI-RADS demonstrated significant discrimination between groups, with disease recurrence/persistence rates of 3.8% for NI-RADS 1, 17.2% for NI-RADS 2, and 59.4% for NI-RADS 3. A strong association between score and positive disease was found for primary site, lymph nodes, and all targets combined, and ROC analysis also demonstrated clinically significant and accurate performance in these categories. While adding additional NI-RADS categories may improve ROC performance, the simplicity of the current scale is appropriate for the limited management options: routine surveillance, shorter interval follow-up, additional PET/CT imaging, mucosal inspection, or biopsy.

Because all these patients are part of our institutional surveillance program, with routine follow-up, it is reasonable for the specificity to be high and sensitivity lower. In fact, size cutoffs were set for the “ill-defined” or “questionable” NI-RADS 2 lesions to avoid low-yield, difficult, dangerous, or likely nondiagnostic

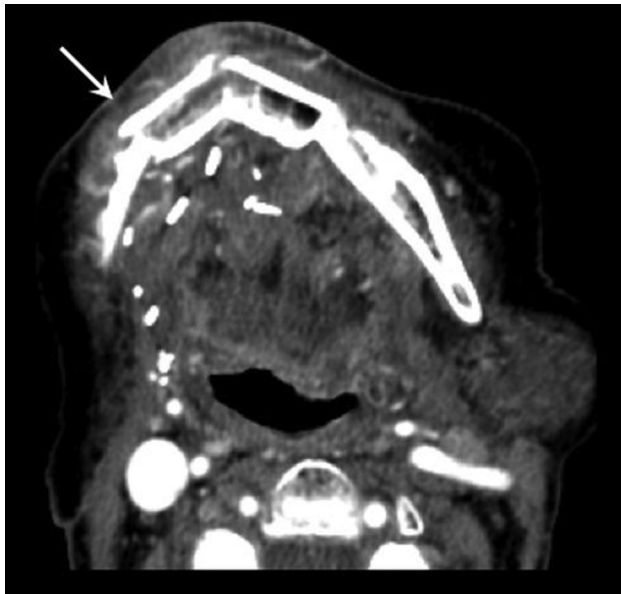


FIG 6. NI-RADS primary site category 2b: ill-defined asymmetric soft tissue. T4N0 oral cavity squamous cell carcinoma. CECT shows asymmetric full soft tissue around fibular reconstruction of the mandible (*arrow*). The linked management recommendation is shorter interval surveillance. Repeat CECT at 3 months showed no interval change (not shown). Subsequent clinical follow-up also demonstrated improvement and no disease recurrence.

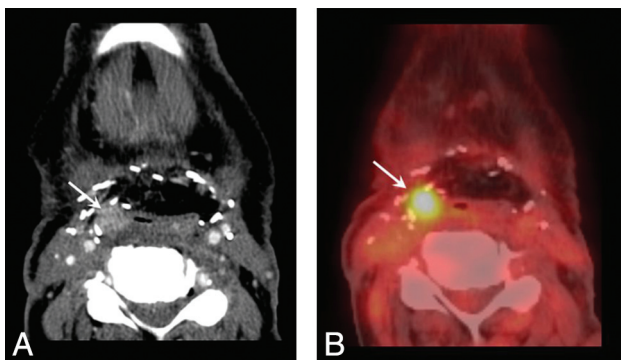


FIG 7. NI-RADS primary site category 3: discrete enhancing lesion. T4a larynx squamous cell carcinoma, status post total laryngectomy, bilateral neck dissection, and chemoradiotherapy. *A*, CECT shows a 1-cm discrete rounded hyperenhancing nodule along the lateral border of neopharynx, deep to the flap (*arrow*). *B*, Fused PET images show focal high FDG uptake (*arrow*). This was given a category 3 score, and endoscopic biopsy demonstrated recurrence.

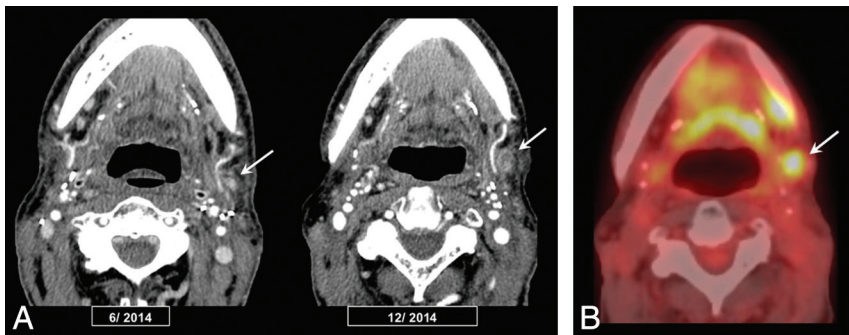


FIG 8. NI-RADS neck category 3: new or enlarged lymph node. T2N0 oral cavity squamous cell carcinoma status post resection, neck dissection, and adjuvant radiation therapy. *A*, CECT at 6-month intervals shows enlarging left level 1B lymph node with necrosis (*arrows*). *B*, Fused PET images show marked focal FDG uptake (*arrow*). Revision neck dissection was positive for disease recurrence.

biopsies for an intermediate suspicion lesion in a complex post-treatment neck if short-interval follow-up or PET would be a viable option. Because size was not helpful in predicting recurrence or for practical management decisions, our revised NI-RADS category 2 does not have size criteria and management consists of earlier follow-up or PET. Biopsy recommendations are reserved for category 3 lesions only. In our experience with highly suspicious lesions, size rarely determines recurrence, compared with enhancement characteristics, morphology, interval change, and FDG uptake, which have been incorporated into the NI-RADS lexicon.

Our NI-RADS template has been useful in daily clinical practice. For the primary site, the 2a category is used for low-suspicion superficial mucosal lesions with a linked recommendation of direct inspection. Focal asymmetric enhancement and FDG uptake are not an uncommon finding in posttreatment imaging and could represent benign mucositis or early recurrence/persistence. Although many mucosal abnormalities are false-positives (Fig 4), we are able to identify mucosal recurrences, especially in the post-radiated larynx where abnormalities may be subtle on CECT (Fig 5A). In this clinical scenario, the fused PET images (Fig 5B) help direct inspection and biopsy. For the primary site, the 2b category is used for deep, ill-defined, nondiscrete, low-suspicion lesions with only mild FDG uptake (if combined with PET) (Fig 6). In practice, most category 2 lesions are managed with short-term follow-up rather than biopsy because clinicians and patients were comfortable waiting. Short-term follow-up has become our official recommendation in this category because size criteria were removed.

Finally, NI-RADS 3 is reserved for a discrete, nodular, robustly enhancing lesion (Fig 7A) with marked FDG uptake if PET was also performed (Fig 7B), and the recommendation is for biopsy. In the neck, NI-RADS 3 is a new or enlarging lymph node (Fig 8A) with marked FDG uptake if PET is combined (Fig 8B). The positive predictive value for NI-RADS 3 primary site lesions was lower (54.6%) than for the neck (70%); this finding likely reflects the more complex posttreatment imaging appearance at the primary site. Overall, we believe that the NI-RADS template yielded a reasonable rate of recommending biopsy. Only 32 of 618 possible targets (5.2%) were scored category 3 with biopsy recommendation, balanced against a relatively high positive predictive value (54.6% for the primary site, 70% for the neck).

Our subgroup analyses highlight areas for future study. Although the numbers are small, our data suggest that CECT alone may be more specific because the rate of true persistence/recurrence was much higher for a NI-RADS 3 for CECT (91.7%) alone versus CECT/PET (40%). We also separated our scans into the first posttreatment baseline at 3 months and subsequent surveillance studies to understand the variation of NI-RADS performance at different time points. As expected, a NI-RADS 1 on a subsequent follow-up examination had a higher negative predictive value than on the initial posttreatment examination. This is valu-

able in providing guidance to patients regarding their risk of disease at different time points.

Finally, NI-RADS provides a meaningful framework for discussion of results with patients. For example, a patient with a NI-RADS 2 on surveillance imaging has a chance of recurrence of roughly 17.2%. We can also reassure patients with NI-RADS 1 that their overall recurrence rate is low (3.8%). There is an opportunity to understand the negative and positive predictive values of NI-RADS scores at different time points in further subpopulation studies. For example, our subgroup analysis and comparison of NI-RADS 1 score primary site recurrence rates for the initial post-treatment examinations versus subsequent examinations found a difference (9.3% versus 2.4%, $P = .047$), but the overall numbers were small because the incidence of recurrence in this group was so low. This analysis suggests that a NI-RADS 1 on the initial baseline posttreatment examination is not as reassuring as a NI-RADS 1 on subsequent surveillance examinations.

CONCLUSIONS

The performance of NI-RADS was good, demonstrating significant discrimination between groups, with positive disease rates of 3.8% for NI-RADS 1, 17.2% for NI-RADS 2, and 59.4% for NI-RADS 3. Standardization of linked management recommendations and correlation with patient outcomes should validate performance and highlight the added value of radiologists in patient care.

REFERENCES

1. Dodd GD 3rd, Allen B Jr, 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
2. Mitchell DG, Bruix J, Sherman M, et al. **LI-RADS (Liver Imaging Reporting and Data System): summary, discussion, and consensus of the LI-RADS Management Working Group and future directions.** *Hepatology* 2015;61:1056–65 CrossRef Medline
3. Weinreb JC, Barentsz JO, Choyke PL, et al. **PI-RADS Prostate Imaging-Reporting and Data System: 2015, Version 2.** *Eur Urol* 2016;69:16–40 CrossRef Medline
4. Cheng SP, Lee JJ, Lin JL, et al. **Characterization of thyroid nodules using the proposed thyroid imaging reporting and data system (TI-RADS).** *Head Neck* 2013;35:541–47 CrossRef Medline
5. Aiken AH, Farley A, Baugnon KL, et al. **Implementation of a novel surveillance template for head and neck cancer: Neck Imaging Reporting and Data System (NI-RADS).** *J Am Coll Radiol* 2016;13:743–46.e1 CrossRef Medline
6. Leung AS, Rath TJ, Hughes MA, et al. **Optimal timing of first post-treatment FDG PET/CT in head and neck squamous cell carcinoma.** *Head Neck* 2016;38(suppl 1):E853–58 CrossRef Medline
7. Gilbert MR, Branstetter BF 4th, Kim S. **Utility of positron-emission tomography/computed tomography imaging in the management of the neck in recurrent laryngeal cancer.** *Laryngoscope* 2012;122:821–25 CrossRef Medline
8. Koshkareva Y, Branstetter BF 4th, Gaughan JP, et al. **Predictive accuracy of first post-treatment PET/CT in HPV-related oropharyngeal squamous cell carcinoma.** *Laryngoscope* 2014;124:1843–47 CrossRef Medline
9. McDermott M, Hughes M, Rath T, et al. **Negative predictive value of surveillance PET/CT in head and neck squamous cell cancer.** *AJNR Am J Neuroradiol* 2013;34:1632–36 CrossRef Medline
10. Pfister DG, Spencer S, Brizel DM, et al; National Comprehensive Cancer Network. **Head and neck cancers, Version 2.2014: clinical practice guidelines in oncology.** *J Natl Compr Canc Netw* 2014;12:1454–87 Medline
11. Roman BR, Patel SG, Wang MB, et al. **Guideline familiarity predicts variation in self-reported use of routine surveillance PET/CT by physicians who treat head and neck cancer.** *J Natl Compr Canc Netw* 2015;13:69–77 Medline
12. Ho AS, Tsao GJ, Chen FW, et al. **Impact of positron emission tomography/computed tomography surveillance at 12 and 24 months for detecting head and neck cancer recurrence.** *Cancer* 2013;119:1349–56 CrossRef Medline
13. Ong SC, Schöde H, Lee NY, et al. **Clinical utility of 18F-FDG PET/CT in assessing the neck after concurrent chemoradiotherapy for locoregional advanced head and neck cancer.** *J Nucl Med* 2008;49:532–40 CrossRef Medline