



## Discover Generics

Cost-Effective CT & MRI Contrast Agents

 FRESENIUS  
KABI

[WATCH VIDEO](#)

# AJNR

This information is current as  
of June 24, 2025.

## **The Impact of Interactive MRI-Based Radiologist Review on Radiotherapy Target Volume Delineation in Head and Neck Cancer**










D. Adjogatse, I. Petkar, M. Reis Ferreira, A. Kong, M. Lei, C. Thomas, S.F. Barrington, C. Dudau, P. Touska, T. Guerrero Urbano and S.E.J. Connor

*AJNR Am J Neuroradiol* 2023, 44 (2) 192-198

doi: <https://doi.org/10.3174/ajnr.A7773>

<http://www.ajnr.org/content/44/2/192>

# The Impact of Interactive MRI-Based Radiologist Review on Radiotherapy Target Volume Delineation in Head and Neck Cancer

 D. Adjogatse,  I. Petkar,  M. Reis Ferreira,  A. Kong,  M. Lei,  C. Thomas,  S.F. Barrington,  C. Dudau,  P. Touska, T. Guerrero Urbano, and  S.E.J. Connor



## ABSTRACT

**BACKGROUND AND PURPOSE:** Peer review of head and neck cancer radiation therapy target volumes by radiologists was introduced in our center to optimize target volume delineation. Our aim was to assess the impact of MR imaging-based radiologist peer review of head and neck radiation therapy gross tumor and nodal volumes, through qualitative and quantitative analysis.

**MATERIALS AND METHODS:** Cases undergoing radical radiation therapy with a coregistered MR imaging, between April 2019 and March 2020, were reviewed. The frequency and nature of volume changes were documented, with major changes classified as per the guidance of The Royal College of Radiologists. Volumetric alignment was assessed using the Dice similarity coefficient, Jaccard index, and Hausdorff distance.

**RESULTS:** Fifty cases were reviewed between April 2019 and March 2020. The median age was 59 years (range, 29–83 years), and 72% were men. Seventy-six percent of gross tumor volumes and 41.5% of gross nodal volumes were altered, with 54.8% of gross tumor volume and 66.6% of gross nodal volume alterations classified as “major.” Undercontouring of soft-tissue involvement and unidentified lymph nodes were predominant reasons for change. Radiologist review significantly altered the size of both the gross tumor volume ( $P = .034$ ) and clinical target tumor volume ( $P = .003$ ), but not gross nodal volume or clinical target nodal volume. The median conformity and surface distance metrics were the following: gross tumor volume Dice similarity coefficient = 0.93 (range, 0.82–0.96), Jaccard index = 0.87 (range, 0.7–0.94), Hausdorff distance = 7.45 mm (range, 5.6–11.7 mm); and gross nodular tumor volume Dice similarity coefficient = 0.95 (0.91–0.97), Jaccard index = 0.91 (0.83–0.95), and Hausdorff distance = 20.7 mm (range, 12.6–41.6). Conformity improved on gross tumor volume-to-clinical target tumor volume expansion (Dice similarity coefficient = 0.93 versus 0.95,  $P = .003$ ).

**CONCLUSIONS:** MR imaging-based radiologist review resulted in major changes to most radiotherapy target volumes and significant changes in volume size of both gross tumor volume and clinical target tumor volume, suggesting that this is a fundamental step in the radiotherapy workflow of patients with head and neck cancer.

**ABBREVIATIONS:** AJCC = American Joint Committee on Cancer; CTV = clinical target volume; CTVN = clinical target nodal volume; CTVT = clinical target tumor volume; DSC = Dice similarity coefficient; GTV = gross target volume; GTVN = gross nodal volume; GTVonc = original oncology GTV; GTVrad = post-radiology review volume; GTVT = gross tumor volume; HD = Hausdorff distance; HN = head and neck; HNC = head and neck cancer; IQR = interquartile range; JI = Jaccard index; PTV = planning target volume; RT = radiotherapy; TRE = target registration error

The efficacy of curative radiotherapy (RT) for head and neck cancers (HNCs) requires the delivery of high doses of radiation to well-defined disease volumes, while sparing normal

tissues as much as possible.<sup>1</sup> HNC target volume delineation relies on accurate interpretation of radiologic imaging with complex locoregional anatomy. While CT provides geometric accuracy and

Received July 14, 2022; accepted after revision December 31.

From the Departments of Oncology (D.A., I.P., M.R.F., A.K., M.L., T.G.U.), Medical Physics (C.T.), and Radiology (C.D., P.T., S.E.J.C.), Guy's and St Thomas' National Health Service Foundation Trust, London, UK; School of Biomedical Engineering and Imaging Sciences (D.A., C.T., S.E.J.C.) and King's College London and Guy's and St Thomas' PET Centre (S.F.B.), School of Biomedical Engineering and Imaging Sciences, King's College London, King's Health Partners, London, UK; Department of Neuroradiology (C.D., S.E.J.C.), King's College Hospital, London, UK; and Faculty of Dentistry, Oral and Craniofacial Sciences (T.G.U.), King's College London, London, UK. S.F. Barrington acknowledges support from the National Institute for Health Research and Social Care (RP-2-16-07-001). King's College London and the UCL Comprehensive Cancer Imaging Center are funded by the Cancer Research UK and EPSRC in association with the Medical Research Council and Department of Health and Social Care (England). This work was also supported by the Wellcome Trust/Engineering and Physical Sciences Research Council Center for Medical Engineering at King's College London (WT 203148/Z/16/Z).

The views expressed are those of the authors and not necessarily those of the National Health Service, the National Institute for Health Research and Social Care, or the Department of Health and Social Care.

For Dr Steve Connor, this research was funded, in whole or in part, by the Wellcome Trust (203148/Z/16/Z). For the purpose of open access, the author has applied a CC BY public copyright license to any author accepted manuscript version arising from this submission. Authors acknowledge funding support from the Wellcome/Engineering and Physical Sciences Research Council Center for Medical Engineering at King's College London (WT 203148/Z/16/Z); the National Institute for Health Research Biomedical Research Center at Guy's and St Thomas' Hospitals and King's College London; the Cancer Research UK National Cancer Imaging Translational Accelerator (A27066); and the UK Research and Innovation London Medical Imaging and Artificial Intelligence Center.

For Teresa Guerrero Urbano, this work was supported by the Radiation Research Unit at the Cancer Research UK City of London Center [C7893/A28990].

relative electron density–derived dose calculations required for RT planning, MR imaging allows more accurate disease identification, owing to improved soft-tissue contrast resolution.<sup>2</sup>

Variability of gross target volume (GTV) delineation has been shown to affect dose distribution to tumors and organs at risk.<sup>3</sup> The uncertainty of GTV delineation is partially accounted for through the creation of radical clinical and planning target volumes (CTV and PTV). For intensity-modulated RT, it is recommended that radical CTVs be created from geometric expansion of GTVs, rather than larger traditional anatomic boundaries.<sup>4</sup> Therefore, precise GTV delineation is crucial for improving plan quality and patient outcomes.<sup>5</sup> MR imaging–based GTV delineation has been shown to reduce interobserver variability.<sup>3</sup> Therefore, coregistration of MR imaging with planning CT scans for RT volume delineation is important for optimum disease definition, with acquisition in the RT immobilization mask and position preferred for superior precision of CT-to-MR imaging registration.<sup>6</sup>

Review of oncologist-defined RT target volumes by head and neck (HN) radiologists reportedly changes 52%–55% of volumes when using CT or diagnostic MR imaging.<sup>7,8</sup> Peer review by fellow oncologists is associated with alteration rates of 14%–39%.<sup>9,10</sup> The importance of oncology peer review is well-established and is the standard of care in many UK centers, whereas radiologist involvement occurs in only 8% of institutions.<sup>11</sup>

In 2018, interactive MR imaging–based radiologist review was formally introduced into the RT workflow for patients with HNC at our institution. It was hypothesized that this would result in significant changes to the RT target volumes. This study aimed to describe the frequency, nature, causes, and clinical significance of alterations and to compare the size and conformity indices for target volumes obtained before and after radiologist review.

## MATERIALS AND METHODS

The study received local institutional approval as a service evaluation (Guy's and St Thomas' NHS Foundation Trust, No. 9623). All patients prospectively consented to use of anonymized information for audit and quality-improvement purposes at the time of their RT consent.

### Participants

From April 2019 to March 2020, consecutive patients undergoing radical volumetric modulated arc therapy for HNC, for which the use of a coregistered MR imaging was indicated for target volume delineation, were prospectively reviewed during weekly interactive oncologist/radiologist sessions.

### Imaging and Coregistration Protocols

RT planning CT scans were acquired on a Biograph mCT Flow (Siemens) at 2.5-mm section thickness with contrast enhancement,

unless contraindicated. All planning CT scans were performed using a 5-point thermoplastic immobilization shell.

Patients had a diagnostic MR imaging performed on either 1.5T or 3T MR imaging Magnetom Aera or Skyra scanners (Siemens) using a surface phased-array 20- or 32-channel neck coil. The diagnostic MR imaging protocol included the following: axial T1WI TSE with and without fat suppression, T2WI TSE, and contrast-enhanced T1WI (plus fat suppression); coronal T2WI STIR; and axial DWI. Section thicknesses ranged from 3 mm (STIR) to 4 mm (anatomic and DWI). Wherever possible, patients with paranasal sinus and nasopharyngeal cancer also underwent MR imaging for RT planning in an immobilization mask. The RT planning MR imaging protocol included the following: sagittal T2WI sampling perfection with application-optimized contrasts by using different flip angle evolution (SPACE sequence; Siemens), axial T1WI FLASH, and contrast-enhanced axial T1WI FLASH (+ fat suppression) sequences, all acquired at 1-mm section thickness.

The MRIs were rigidly coregistered with the planning CT within the RT treatment planning system. The institutional protocol in use during the study period mandated the diagnostic or RT planning T2WI sequence for coregistration, because historically this sequence had most frequently been referenced for volume delineation. Additional sequences were coregistered on request or at the discretion of the RT pretreatment team. Coregistration accuracy was retrospectively determined within the RT treatment planning system using the target registration error (TRE) of 1 bony and 2 soft-tissue landmarks. The mean (TRE<sub>mean</sub>) and maximum (TRE<sub>max</sub>) values for each coregistration were obtained and assessed against the recommended optimum of TRE<sub>mean</sub> of  $\leq 2$  mm and TRE<sub>max</sub> of  $\leq 5$  mm.<sup>12</sup>


### Radiologist Review and Definition of Target Volumes


Gross tumor and gross tumor nodal volumes (GTVTs and GTVNs) were generated by 1 of 5 clinical oncologists (with 2, 2, 7, and 12 years of HNC consultant experience) on Eclipse RT Treatment Planning System (Version 15.5; Varian), with concurrent use of Sectra PACS IDS7 (Sectra) and supportive clinical information. Volumes were subsequently reviewed jointly with 1 of 3 HN radiologists (with 2, 6, and 20 years' experience), also using Eclipse and PACS. The review process included interactive discussion with reference to clinical information and other diagnostic imaging (eg, PET/CT) at the discretion of the radiologist. All radiologists and oncologists were aware of the study purpose.

The primary focus was review of the delineation of all gross disease on MR imaging. GTVs were duplicated and saved before radiology review to preserve the original oncology GTV (GTV<sub>onc</sub>). A second postradiology review volume (GTV<sub>rad</sub>) was created and amended as necessary. Contours were adjusted by the radiologist or the oncologist under the instruction of the radiologist. All final amended GTVs were also viewed on the CT planning scan to assess discrepancies, which were noted and amended when possible. Patients proceeded to the creation of CTVs and PTVs as per institutional protocol, with no further input from radiologists.

For this evaluation, study CTVs were created by a single observer (the first author of this report who is a clinical oncologist not part of the original delineation or peer review process) to avoid interobserver variability in the CTV delineation process and to

Please address correspondence to Delali Adjogatse, MD, Guy's and St Thomas' Hospitals NHS Foundation Trust, Guy's Hospital, Great Maze Pond, SE1 9RT; e-mail: Delali.adjogatse@gstt.nhs.uk

 Indicates open access to non-subscribers at [www.ajnr.org](http://www.ajnr.org)

 Indicates article with online supplemental data.

<http://dx.doi.org/10.3174/ajnr.A7773>

assess the impact of radiology review on CTVs. This process involved geometric expansion of both GTVonc/GTVrad by a 10-mm isotropic margin to create tumor or nodal CTVonc/CTVrad (20 mm for nodes with evidence of extranodal spread) and editing off barriers to tumor spread (ie, air, bone, and muscle if muscle invasion was absent). In postoperative cases, the preoperative disease on MR imaging was used to guide the extent of the CTV, which would ultimately encompass the entire anatomic surgical bed. Therefore, adjustments to preoperative GTVs were included in this study, but not the postoperative CTVs.

### **Radiation Therapy Planning**

All patients underwent volumetric modulated arc therapy, with standard PTV doses and fractionations as follows: 65 Gy in 30 fractions (f) to cover gross disease; 60 Gy/30f to the postoperative tumor bed; and 54 Gy/30f to elective nodal regions, delineated as per consensus guidelines.<sup>13,14</sup> Isotropic 4-mm margins were added to CTVs to create radical PTVs.

### **Descriptive Data and Qualitative Analysis**

Patient demographics, tumor subsite and staging (American Joint Committee on Cancer [AJCC] and tumor, node and metastasis), delineating oncologist, attending radiologist, and time taken for review were documented prospectively. The nature, anatomic patterns, and contributing factors for volume amendment were noted at the time of review. Anatomic patterns were the following: lymph node (addition/removal), GTVT deep extent (changes to submucosal involvement), GTVT superficial extent (changes primarily along mucosal surface), normal structure exclusion, skull base/bone or perineural spread, intracranial, and sinonasal (intra-sinus) extension. Oncologists also prospectively recorded whether they thought the changes made at peer review were clinically significant. This judgment was qualitative, combining a subjective assessment of clinical significance and a predetermined definition of geographic miss (Online Supplemental Data).

### **Volumetric and Quantitative Analysis**

The absolute and percentage volume differences between the GTVonc and GTVrad (regarded as the expert volume) were recorded. Conformity (degree of spatial overlap) was assessed using the Dice similarity coefficient (DSC) and Jaccard index (JI).<sup>15</sup> The Hausdorff distance (HD) assessed the maximum distance between voxel locations for each tumor volume, to assess the deviation between volumes in 3D.<sup>16</sup> These indices were obtained using MIM Encore (MIM Software). Only amended volumes that were duplicated and saved before peer review were included in volumetric analysis.

Postradiology volume adjustments were retrospectively classified as major or minor on the basis of the guidance of the Royal College of Radiologists (Online Supplemental Data).<sup>17</sup> Alterations to prevent a geographic miss, such as editing a GTVT by  $\geq 10$  mm or adding suspicious nodes within the radical CTV, were defined as major. Alterations that would otherwise still be clinically acceptable, such as editing of normal tissues, were classified as minor.

### **Statistical Analysis**

Descriptive statistics documented the frequency of GTV change, anatomic patterns of change, and rates of major change. The

Shapiro-Wilk test was used to assess data-distribution normality. The Spearman correlation coefficient assessed an association between years of experience (oncologist/radiologist) and the frequency of any change and major change. The difference in volume size between pre- and postradiology GTVs and CTVs and the differences in the distribution of DSC, JI, and HD between GTVs and CTVs were analyzed using the Wilcoxon signed-rank test. A 2-tailed Fisher exact test was used to determine whether categorical factors were associated with any or major change. Analyzed factors were the following: histology, T stage, N stage, AJCC stage, postoperative status, use of a research trial delineation protocol, and use of DWI for target identification. Registration error (TREmean and TREmax) and its correlation with frequency of any and major change were assessed using the Mann-Whitney *U* test. Statistical significance was  $P < .05$ .

## **RESULTS**

### **Participant Characteristics**

We reviewed 50 consecutive patients (Online Supplemental Data): 37 patients with definitive intensity-modulated RT, 11 postoperative patients, and 2 high-dose palliative patients. The median age was 59 years, 72% were men, and 72% had oropharyngeal cancer. There were 50 GTVTs and 42 GTVNs reviewed. One GTVN review was subsequently excluded as new diagnostic information became available (fine-needle aspiration cytology) after GTVNonc delineation. The median radiology time taken per case was 20 minutes (interquartile range [IQR] = 15–30; absolute range = 5–60). Five patients had MR imaging scans for RT planning; all others had diagnostic MRIs. Figure 1 illustrates an example of post-peer review volume adjustments.

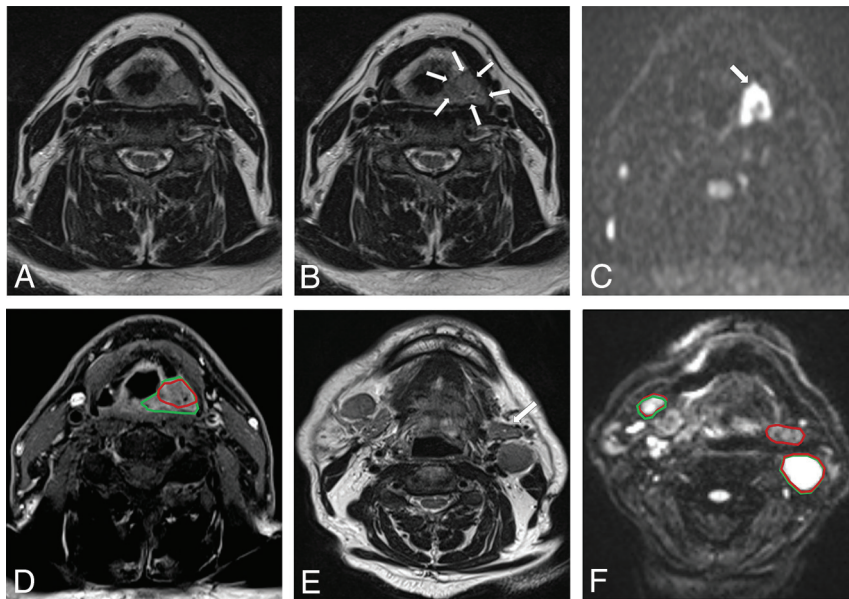
### **Frequency, Nature, and Anatomic Patterns of GTV and CTV Changes following MR Imaging–Based Radiologist Review**

Forty-two patients had at least 1 GTV amended (84%). Seventy-six percent of GTVTs (38/50) and 41.5% of GTVNs (17/41) were amended.

Anatomic patterns of change were the following: GTVN, additional nodes identified (82.3%; 14/17); GTVT, deep extent (68.4%; 26/38); GTVT, superficial extent (23.6%, 9/38); normal structure exclusion (GTVT, 5.2%, 2/38; GTVN, 11.8%, 2/17); and GTVT with perineural spread (2.5%, 1/38).

Explanatory documentation for GTV modification was available for all amended volumes and was retrospectively organized into 5 distinct groups (Fig 2). The most common reasons for modifications were imaging misinterpretation (GTVT, 57.8%, 22/38; GTVN, 52.9%, 9/17) and changes made after collaborative discussion (GTVT, 7.9%, 3/38; GTVN, 23.5%, 4/17). Medial extension of oro-/hypopharyngeal tumors ( $n = 6$ ) and superior extension of nasopharyngeal tumors into the nasopharyngeal vault ( $n = 4$ ) were common causes for GTVT alteration. The most frequent reasons for GTVN addition were inclusion of suspicious or pathologic nodes adjacent to correctly delineated nodal disease in cervical levels 1b–3 ( $n = 9$ ) and inclusion of retropharyngeal nodes ( $n = 3$ : 1 missed despite diagnostic report, 1 pathologic but unreported, 1 highly suspicious and unreported). Of all patients with amended volumes,





**FIG 1.** Examples of volume amendments made at peer review. Selected images show the difference in GTVonc and GTVrad delineations. A and B, T2-weighted axial images show an intermediate-signal left piriform fossa tumor (arrows). Following review of the diffusion-weighted imaging ( $b = 1000$ ) (single arrow) (C), the contouring on T1-weighted gadolinium-enhanced axial image is expanded from GTVonc (red) to GTVrad (green) (D). T2-weighted axial image demonstrates multiple bilateral lymph nodes. The left submandibular gland (arrow) was initially included in the GTVonc (red) because it was isointense to other pathologic lymph nodes. F, Diffusion-weighted imaging ( $b = 1000$ ) aided in the identification of the lower signal submandibular gland; hence, it was excluded from GTVrad (green).

Imaging misinterpretation	Clinician Discretion <sup>1</sup>	Use of additional DWI sequences <sup>2</sup>	Complex case <sup>3</sup>	Use of additional imaging (eg, PET)
Volume added 22	Volume added 7	Volume added 7	Volume added 3	Volume added 3
Volume removed 9	Volume removed 0	Volume removed 0	Volume removed 1	Volume removed 1

**FIG 2.** Factors contributing to volume amendments: GTVT and GTVN. <sup>1</sup>Changes made following joint discussion between the oncologist and radiologist, eg, inclusion of lymph nodes with borderline pathologic changes. <sup>2</sup>Diffusion-weighted MR imaging sequences. <sup>3</sup>Tumor volumes described as complex by the oncologist and radiologist at the time of peer review. Includes skull base disease with perineural spread, septate tumor, and postexcisional biopsy changes.

76.2% (32/42) were judged as having clinically significant alterations by the delineating oncologist.

AJCC stage 4 was the only covariate associated with any change to the GTVT ( $P = .043$ , OR = 5.33; 95% CI, 1.078–26.35). Factors associated with any GTVN amendment were the following: nodal stages 1 ( $P = .029$ , OR = 0.77; 95% CI, 0.48–0.97) and 3 ( $P = .008$ , OR = 1.42; 95% CI, 1.04–1.93) and AJCC stages 3 ( $P = .028$ , OR = 0.10; 95% CI, 0.01–0.92) and 4 ( $P = .014$ , OR = 11.43; 95% CI, 1.29–100.82).

### Volumetric and Quantitative Analysis

Of the 55 amended volumes, 40 were available for volumetric analysis: 31 GTVTs and 9 GTVNs. Seven GTVTs and 8 GTVNs

were excluded because the original GTVonc was not preserved. In addition, 22 study clinical target tumor volumes (CTVTs) and 9 study clinical target nodular volumes (CTVNs) were analyzed. Nine postoperative CTVTs were excluded because changes to the GTVT did not affect the CTVT.

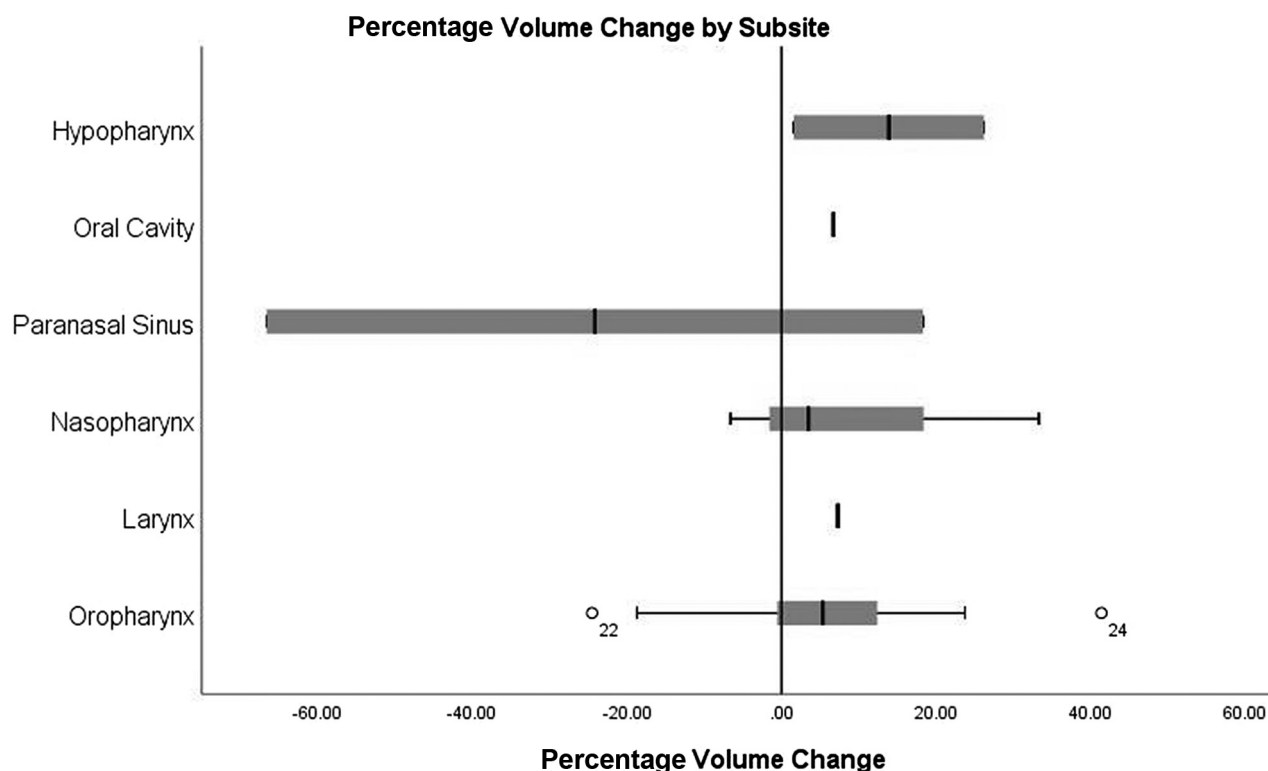
Volumetric results of amended volumes are detailed in the Online Supplemental Data. Of these, major changes were seen in 54.8% of GTVTs (17/31) and 66.6% of GTVNs (6/9). Seventy-four percent of GTVTs (23/31) and 77.7% of GTVNs (7/9) were increased. The median percentage volume change was 5.7% and 4.5% for GTVT and GTVN, respectively. An increase in volume was seen across most subsites (Fig 3). Overall volume similarity was very good for GTVT (DSC = 0.93, JI = 0.87, HD = 7.45 mm) and GTVN (DSC = 0.95, JI = 0.91, HD = 20.7 mm). The similarity of tumor volumes increased after GTVT-to-CTVT expansion. Following GTVN-to-CTVN expansion volume similarity varied; surface distance (HD) reduced, however volume overlap (DSC/JI) worsened.

There was a statistically significant difference between the volume sizes of GTVTonc and GTVTrad ( $P = .034$ ) and CTVTonc and CTVTrad ( $P = .003$ ), but not for GTVN or CTVN. Conformity was significantly higher for CTVT versus GTVT with a median DSC of 0.95 versus 0.93 ( $P = .003$ ) and a median JI of 0.90 versus 0.87 ( $P = .003$ ). There were no statistically significant differences in conformity between CTVN and GTVN.

There were no covariates associated with major changes to either GTVT or GTVN. Increasing years of oncologists' experience correlated with increased DSC and JI ( $P = .018$ ) for GTVT. There was no correlation for radiologists' years of experience. Rates of any and major change did not differ depending on the clinician's experience.

### Coregistration Error

Fifty-five registration errors were assessed (5 RT planning, 50 diagnostic scans). The median TREmean and TREmax for all coregistrations was 4.5 mm (IQR = 3.15–5.6) and 6.4 mm (IQR = 4.6–8.45), respectively. For RT planning scans, the median TREmean was 2.2 mm (IQR = 1.65–2.75), and the TREmax, 2.8 mm (IQR = 1.9–3.7). Registration error exceeded both optimum TRE values in 52.7% ( $n = 29$ ) of scans.



**FIG 3.** Boxplots of percentage volume change for GTVT by subsite. Case number 22 and 24 (highlighted as superscript numbers on the box plot of oropharyngeal cancer cases) are outliers in terms of percentage volume change. Case 22: the GTVT was reduced by 24.5% on radiologist review due to the removal of normal oral tongue and normal parapharyngeal fat. Case 24; GTVT increased by 41.4% on radiologist review due to imaging misinterpretation of abnormal deep mucosal disease extent.

Registration error did not correlate with the frequency of any change; however, the distribution of TREmean values was higher for cases that underwent major-versus-minor change: median, 4.9 mm (IQR = 3.23–6.53) versus 3.6 mm (IQR = 1.97–5.23), respectively ( $P = .048$ ). No differences were seen in the distribution of TREmax.

## DISCUSSION

MR imaging–based radiologist review resulted in amendments to 76% of tumor and 41.4% of nodal volumes, with 54.8% of GTVT and 66.6% of GTVN changes being classified as major. Most GTVT changes were related to the addition of disease in the deep mucosal extent. Some decisions made following collaborative discussion, such as inclusion of borderline nodes, may not have occurred during oncologist-only contouring, and their clinical relevance is uncertain, with a potential for overcontouring resulting in an increased dose to normal tissues. Conformity indices suggested good similarity between the pre- and postradiologist GTVT and GTVN. Because expansion from GTVT was often constrained by adjacent bone and air, there was less impact of the radiologist's review on representative CTVT volumes, though the small differences in volume size were statistically significant. Conversely, differences among nodal volumes were amplified by expansion to the CTVN. This did not reach statistical significance, possibly due to the small number of nodal volumes available for analysis.

Among radiologists, HN imaging is recognized as a challenging subject requiring specialized training.<sup>18</sup> Oncology training

includes no formal diagnostic radiology education; therefore, radiologist-delineated volumes were presumed to be the criterion standard. In our study, most GTV changes were due to undercontouring. Radiologists routinely analyzed information from multiple sequences and had more familiarity with additional sequences such as DWI, which allows greater confidence in distinguishing pathology from normal tissues and in detecting lymph nodes. Review of the DWI, in particular, was implicated in 7 GTV adjustments on peer review. Oncologists agreed with radiologists' recommendations to increase volumes to reduce the risk of a geographic miss, as primary tumor recurrence within the irradiated field is rarely salvageable. However, if the radiologic evaluation of superficial tumor extension was thought to be erroneous on the basis of mucosal clinical findings, then this issue was discussed and discounted.

Diagnostic reports were available at contouring; however, most lymph node additions were either misidentified pathologic nodes within the same cervical level as correctly delineated nodes or suspicious nodes that were not formally reported. In our institution, diagnostic reports state the existence of nodal disease per cervical nodal level without listing each individual node, thus increasing the likelihood of imaging misinterpretation by oncologists. In addition to detecting overlooked pathologically appearing (by size and morphological criteria) lymph nodes, there was a tendency for radiologists to suggest the inclusion of borderline enlarged nodes in the lymphatic drainage pathways of the primary tumor that were deemed suspicious, with which the oncologists agreed.

Compared with similar studies, our rate of change to gross volumes was high. However, the proportion of major changes was comparable. Braunstein et al<sup>7</sup> studied the impact of neuroradiologist peer review on, primarily, CT-based RT planning volumes. They noted a 55% alteration rate, 61% of which was clinically significant. More recently, Chiu et al<sup>8</sup> evaluated the effects of radiologist input using coregistered MR imaging scans and reported a 52% alteration rate, with major changes seen in 79%. Studies of oncologist-only peer review reported lower rates of any change (14%–39%) and major/clinically significant change (8.8%–13%).<sup>9,10</sup> The discrepancy between outcomes of radiologist and oncologist peer review may be reflective of the level of expertise in imaging interpretation, possibly exaggerated further when using MR imaging. Moreover, in our institution, radiologists attended interactive review sessions rather than postdelineation peer review meetings. This practice allowed detailed exploration for each case on a one-to-one basis.

Comparison of clinical significance across studies is difficult because definitions for major/clinically significant changes vary. Some articles, such as this one, classified major changes on the basis of volumetric results; therefore, rates were reported only among amended volumes (with prereview volumes saved).<sup>7</sup> Others reported clinical significance among all reviewed volumes (including unaltered and unsaved volumes).<sup>8–10</sup> Among all reviewed volumes in this study, 34% of GTVT and 14.6% of GTVN would have undergone major changes. However, the lack of volumetric analysis of 15 unsaved cases of GTVonc likely distorts those results.

Our conformity indices were very good and consistent with quality-assurance studies in clinical trials reporting a median JI of  $\geq 0.7$  and a DSC of  $\geq 0.8$ ,<sup>19,20</sup> and the recent peer-review study of Chiu et al.<sup>8</sup> Chiu et al reported median GTVT values of DSC = 0.97, JI = 0.94, and HD = 3.6 mm. They also observed a larger HD for GTVN at 37 mm and a reduction in conformity on GTVN-to-CTVN expansion. While a GTVT is usually 1 continuous volume, GTVN often includes multiple separate nodes; therefore, the inclusion of additional nodes will have a greater impact on conformity and distance. In our study, nodes were often added superior and/or inferior to the GTVNonc, expanding the volume within the soft tissues before anatomic barriers were encountered.

Following publication of the study of Braunstein et al,<sup>7</sup> the use of MR imaging-guided RT volume delineation in HNC has increased, with a projected future move toward delineation based on functional imaging and MR imaging-based synthetic CTs,<sup>21</sup> hence the focus of our study on MR imaging-based volume delineation. In contrast to earlier, predominantly CT-based studies, we detected a predominance of undercontoured volumes by oncologists when using MR imaging.<sup>7,22</sup> This may reflect the lack of formal training in MR imaging interpretation for oncologists as well as a conservative approach to include areas/nodes considered suspicious to minimize the risk of a geographic miss. Therefore, involvement of input of specialist radiologists in GTV delineation may be more advantageous, given the increasing complexity of HN MR imaging interpretation, particularly when using functional sequences.

Few MRIs in our study were acquired in the RT treatment position; therefore, registration error exceeded both optimum

TRE values in half of the scans. The effect of suboptimal registration on the degree of volume change is uncertain. Among amended volumes with TREmax > 5mm ( $n = 24$ ), only 1 case corrected the GTVrad, but not GTVonc, on CT. In 3 cases, both GTVonc/rad were contoured on both CT/MR imaging and compared accordingly. In 5 cases, both GTVonc and GTVrad were produced on MR imaging only. In 16 cases, GTVonc and GTVrad were evidently corrected for anatomic mismatch to some degree (ie, both contours edited away from bone/air). In these cases, some soft-tissue discrepancies remained, which may have been due to registration error, but there was no clarifying documentation regarding this aspect. Nevertheless, the accompanying descriptive reasons for volume change for each case suggest that many soft-tissue volume discrepancies would have persisted irrespective of registration error.

There are limitations to this study. It describes a one-to-one comparison for a single oncologist and single radiologist at a given time, without analysis of interobserver agreement within the oncologist and radiologist groups. It is recognized that delineation of GTVs by each of the oncologists and radiologists would have allowed the calculation of agreement parameters for tumor volumes and would have given a better understanding of performance between and across the 2 disciplines. However, the inclusion of multiple different radiologists and oncologists in the review process reflects the range of real-world practice in our institution.

Another limitation includes a lack of longitudinal data to assess the presence of a learning curve, especially important as the more junior oncologists gain experience. The possibility that oncologists may have delayed complex contouring decisions in anticipation of the pending radiologist review, resulting in more volume amendments, cannot be excluded. However, the prospective documentation suggests that when oncologists viewed a case as complex, the radiologist tended to agree. Therefore, the peer review process facilitated decision-making for these difficult cases.

The use of a single observer to create study CTVs may have increased the variation between the GTVs and CTVs. Ideally the oncologist for each case would have produced a CTVonc before peer review; however, this step was impractical because radiologist review and oncology target delineation sessions occurred on the same day. Preoperative GTVTs were included in volumetric analysis because they still provided information on delineation accuracy of the oncologist. However, removing the consequent postoperative CTVTs may have introduced selection bias toward poorer CTVT conformity (no postoperative nodal volumes were amended). Nevertheless, CTVT conformity was still superior to GTVT conformity on both measures.

Additional comparison metrics assessing the extent of undercontouring (geographic miss index) and overcontouring (discordance index) may have provided a more specific assessment of contour similarity, uninfluenced by volume size. However, further metrics would not likely reduce the qualitative benefits of peer review highlighted in this study. Furthermore, as more centers adopt the intensity-modulated RT 5 + 5 consensus,<sup>23</sup> which reduces the margin for the high-risk CTV to 5mm, the need for optimized GTV delineation is increasing.

At our institution, HN oncology and radiology teams agreed on the need for expert radiology review of gross volumes when



MR imaging delineation was introduced in the HNC RT workflow, given the complexity of the tumor site and the lack of specific MR imaging training of oncologists. This study was planned to assess its impact. Since its implementation, we have expanded the use of MRIs in the immobilization mask for patients having definitive intensity-modulated RT, to reduce registration errors. The efficiency and responsiveness of the radiology review process have been enhanced since 2020 by performing remote radiology review using an application equipped with shared screen functions. CTV and PTV volumes are subsequently peer reviewed within the oncology group. In this study, we have shown that the impact of radiology review of tumor and nodal volumes is sufficiently significant to warrant expert review of radiologists in MR imaging-based volume delineation in patients with HNC, something to be considered by institutions and funding bodies.

## CONCLUSIONS

Interactive MR imaging-based radiologist review resulted in RT target volume amendments in most patients. Despite high conformity indices, most changes to GTVT and GTVN were considered major. Although volumetric similarities improved on GTVT-to-CTVT expansion, changes to the volume size remained statistically significant for both GTVT and CTVT. The true clinical significance of these changes remains uncertain. However, the quantitative measures and descriptive reasons for volume adjustment illustrate the benefit of the input of radiologists for MR imaging-based volume delineation in HNC.

**Disclosure forms** provided by the authors are available with the full text and PDF of this article at [www.ajnr.org](http://www.ajnr.org).

## REFERENCES

- Jackson A, Marks LB, Bentzen SM, et al. **The lessons of QUANTEC: recommendations for reporting and gathering data on dose-volume dependencies of treatment outcome.** *Int J Radiat Oncol Biol Phys* 2010;76:S155–60 [CrossRef Medline](#)
- Chang AE, Matory YL, Dwyer AJ, et al. **Magnetic resonance imaging versus computed tomography in the evaluation of soft tissue tumors of the extremities.** *Ann Surg* 1987;205:340–48 [CrossRef Medline](#)
- Bird D, Scarsbrook AF, Sykes J, et al. **Multimodality imaging with CT, MR and FDG-PET for radiotherapy target volume delineation in oropharyngeal squamous cell carcinoma.** *BMC Cancer* 2015;15:844 [CrossRef Medline](#)
- Head and Neck Cancer: RCR Consensus Statements.** The Royal College of Radiologists; 2022. <https://www.rcr.ac.uk/publication/head-and-neck-cancer-rcr-consensus-statements>. Accessed November 1, 2022
- Van Herk M. **Errors and margins in radiotherapy.** *Semin Radiat Oncol* 2004;14:52–64 [CrossRef Medline](#)
- Prestwich RJ, Sykes J, Carey B, et al. **Improving target definition for head and neck radiotherapy: a place for magnetic resonance imaging and 18-fluoride fluorodeoxyglucose positron emission tomography?** *Clin Oncol (R Coll Radiol)* 2012;24:577–89 [CrossRef Medline](#)
- Braunstein S, Glastonbury CM, Chen J, et al. **Impact of neuroradiology-based peer review on head and neck radiotherapy target delineation.** *AJNR Am J Neuroradiol* 2017;38:146–53 [CrossRef Medline](#)
- Chiu K, Hoskin P, Gupta A, et al. **The quantitative impact of joint peer review with a specialist radiologist in head and neck cancer radiotherapy planning.** *Br J Radiol* 2022;95:20211219 [CrossRef Medline](#)
- Fong C, Sanghera P, Good J, et al. **Implementing head and neck contouring peer review without pathway delay: the on-demand approach.** *Clin Oncol (R Coll Radiol)* 2017;29:841–47 [CrossRef Medline](#)
- Ramasamy S, Murray LJ, Cardale K, et al. **Quality assurance peer review of head and neck contours in a large cancer centre via a weekly meeting approach.** *Clin Oncol (R Coll Radiol)* 2019;31:344–51 [CrossRef Medline](#)
- Mercieca S, Belderbos J, Gilson D, et al. **Implementing the royal college of radiologists' radiotherapy target volume definition and peer review guidelines: more still to do?** *Clin Oncol (R Coll Radiol)* 2019;31:706–10 [CrossRef Medline](#)
- Brock KK, Mutic S, McNutt TR, et al. **Use of image registration and fusion algorithms and techniques in radiotherapy: report of the AAPM Radiation Therapy Committee Task Group No. 132.** *Med Phys* 2017;44:e43–76 [CrossRef Medline](#)
- Grégoire V, Eisbruch A, Hamoir M, et al. **Proposal for the delineation of the nodal CTV in the node-positive and the post-operative neck.** *Radiother Oncol* 2006;79:15–20 [CrossRef Medline](#)
- Grégoire V, Levendag P, Ang KK, et al. **CT-based delineation of lymph node levels and related CTVs in the node-negative neck: DAHANCA, EORTC, GORTEC, NCIC, RTOG consensus guidelines.** *Radiother Oncol* 2003;69:227–36 [CrossRef Medline](#)
- Eelbode T, Bertels J, Berman M, et al. **Optimization for medical image segmentation: theory and practice when evaluating with Dice score or Jaccard index.** *IEEE Trans Med Imaging* 2020;39:3679–90 [CrossRef Medline](#)
- Buttenlocher DP, Klanderman GA, Rucklidge WJ. **Comparing images using the Hausdorff distance.** *IEEE Trans Pattern Anal Mach Intell* 1993;15:850–63 [CrossRef](#)
- Royal College of Radiologists. Radiotherapy Target Volume Definition and Peer Review—RCR guidance.** <https://www.rcr.ac.uk/publication/radiotherapy-target-volume-definition-and-peer-review-second-edition-rcr-guidance>. Accessed December 2021
- Lysack JT, Hoy M, Hudon ME, et al. **Impact of neuroradiologist second opinion on staging and management of head and neck cancer.** *J Otolaryngol Head Neck Surg* 2013;42:39 [CrossRef Medline](#)
- Gwynne S, Spezi E, Wills L, et al. **Toward semi-automated assessment of target volume delineation in radiotherapy trials: the SCOPE 1 pre-trial test case.** *Int J Radiat Oncol Biol Phys* 2012;84:1037–42 [CrossRef Medline](#)
- Tsang Y, Hoskin P, Spezi E, et al. **Assessment of contour variability in target volumes and organs at risk in lung cancer radiotherapy.** *Tech Innov Patient Support Radiat Oncol* 2019;10:8–12 [CrossRef Medline](#)
- Wong KH, Panek R, Bhide SA, et al. **The emerging potential of magnetic resonance imaging in personalizing radiotherapy for head and neck cancer: an oncologist's perspective.** *Br J Radiol* 2017;90:20160768 [CrossRef Medline](#)
- Giraud P, Elles S, Helfre S, et al. **Conformal radiotherapy for lung cancer: different delineation of the gross tumor volume (GTV) by radiologists and radiation oncologists.** *Radiother Oncol* 2002;62:27–36 [CrossRef Medline](#)
- Grégoire V, Evans M, Le QT, et al. **Delineation of the primary tumour clinical target volumes (CTV-P) in laryngeal, hypopharyngeal, oropharyngeal and oral cavity squamous cell carcinoma: AIRO, CACA, DAHANCA, EORTC, GEORCC, GORTEC, HKNPCSG, HNCIG, IAG-KHT, LPRHHT, NCIC CTG, NCRI, NRG ONCOLOGY, PHNS, SBRT, SOMERA, SRO, SSHNO, TROG consensus guidelines.** *Radiother Oncol* 2018;126:3–24 [CrossRef Medline](#)