Brief Communication

Yonsei Med J 2020 Oct;61(10):895-900
https://doi.org/10.3349/ymj.2020.61.10.895

Yonsei Medical Journal
YMJ
pISSN: 0513-5796 · eISSN: 1976-2437

Squamous Cell Carcinoma and Lymphoma of the Oropharynx: Differentiation Using a Radiomics Approach

Sohi Bae[1], Yoon Seong Choi[2], Beomseok Sohn[2], Sung Soo Ahn[2], Seung-Koo Lee[2], Jaemoon Yang[2], and Jinna Kim[2]

[1]Department of Radiology, National Health Insurance Service Ilsan Hospital, Goyang;
[2]Department of Radiology and Research Institute of Radiological Science, Yonsei University College of Medicine, Seoul, Korea.

The purpose of this study was to evaluate the diagnostic performance of magnetic resonance (MR) radiomics-based machine learning algorithms in differentiating squamous cell carcinoma (SCC) from lymphoma in the oropharynx. MR images from 87 patients with oropharyngeal SCC (n=68) and lymphoma (n=19) were reviewed retrospectively. Tumors were semi-automatically segmented on contrast-enhanced T1-weighted images registered to T2-weighted images, and radiomic features (n=202) were extracted from contrast-enhanced T1- and T2-weighted images. The radiomics classifier was built using elastic-net regularized generalized linear model analyses with nested five-fold cross-validation. The diagnostic abilities of the radiomics classifier and visual assessment by two head and neck radiologists were evaluated using receiver operating characteristic (ROC) analyses for distinguishing SCC from lymphoma. Nineteen radiomics features were selected at least twice during the five-fold cross-validation. The mean area under the ROC curve (AUC) of the radiomics classifier was 0.750 [95% confidence interval (CI), 0.613–0.887], with a sensitivity of 84.2%, specificity of 60.3%, and an accuracy of 65.5%. Two human readers yielded AUCs of 0.613 (95% CI, 0.467–0.759) and 0.663 (95% CI, 0.531–0.795), respectively. The radiomics-based machine learning model can be useful for differentiating SCC from lymphoma of the oropharynx.

Key Words: Radiomics, oropharynx, squamous cell carcinoma, lymphoma, magnetic resonance imaging

Squamous cell carcinoma (SCC) is the most common primary malignancy of the oropharynx, followed by lymphoma.[1] Although pretreatment imaging evaluation of these two malignant tumors is critical for tumor staging and treatment planning, it is not easy to differentiate these tumors on the basis of conventional magnetic resonance (MR) imaging alone, because both SCC and lymphoma of the oropharynx usually have similar imaging characteristics.[2]

Previous studies have reported that advanced MR imaging techniques, such as diffusion-weighted imaging[2-5] and dynamic contrast-enhanced MR imaging,[6-8] may be helpful in the differentiation of head and neck cancer. However, these MR imaging sequences are time consuming and may not be performed routinely in daily practice. However, conventional MR imaging with T2-weighted images (T2WIs) and contrast-enhanced T1-weighted images (CE-T1WIs) are almost always available.

A recently introduced radiomics model, which is based on the extraction of high-dimensional quantitative imaging features, can potentially provide underlying pathophysiological information that is difficult to ascertain during visual inspection.[9,10] Accordingly, several previous studies have applied radiomics on head and neck cancers:[11-13] One of these studies found that seven texture features predicted tumor suppressor protein p53 status of oropharyngeal SCC with an accuracy of 81.3%.[11] Another study showed that radiomic features extract-

Received: April 20, 2020 **Revised:** July 31, 2020
Accepted: August 17, 2020
Corresponding author: Jinna Kim, MD, PhD, Department of Radiology and Research Institute of Radiological Science, Yonsei University College of Medicine, 50-1 Yonsei-ro, Seodaemun-gu, Seoul 03722, Korea.
Tel: 82-2-2228-2392, Fax: 82-2-2227-8337, E-mail: jinna@yuhs.ac

•The authors have no potential conflicts of interest to disclose.

ed from square-shaped regions-of-interest in sinonasal tumors could differentiate SCC from inverted papilloma with an accuracy of 89.1%.[12] The purpose of this study was to evaluate the diagnostic performance of MR radiomics-based machine learning algorithms for the differentiation of SCC from lymphoma in the oropharynx.

This single-center retrospective study was approved by the Institutional Review Board of Severance Hospital (IRB no.4-2018-0761), and the requirement for informed consent was waived. The records of 97 consecutive patients with a new diagnosis of SCC or lymphoma of the oropharynx who underwent preoperative MR imaging from July 2012 to June 2018 were reviewed retrospectively. Patients were included on the basis of the following criteria: 1) pathologically confirmed SCC or lymphoma of the oropharynx; 2) preoperative MR imaging with identical sequence protocols, including T2WIs and three-dimensional (3D) fat-saturated CE-T1WIs. The exclusion criteria were as follows: 1) a history of biopsy or treatment before MR imaging (n=7) and 2) inadequate image quality (n=3). Thus, a total of 87 patients were included in this study (Fig. 1). The study group consisted of 68 SCC (58 men and 10 women; mean age±standard deviation, 59.8±9.9 years) and 19 lymphoma patients (14 men and 5 women; 59.9±14.0 years). The primary tumor sites of 68 SCC patients were the palatine tonsil (n=61), base of the tongue (n=6), and soft palate (n=1).

Final histopathological diagnoses were confirmed by fiberscope-guided biopsy within a period of 2 weeks following MR imaging, before the initiation of treatment. The SCCs comprised 15 well-differentiated, 32 moderately differentiated, and 21 poorly differentiated carcinomas. Lymphomas consisted of 17 diffuse large B-cell lymphomas and two Burkitt's lymphomas.

Preoperative MR imaging was performed at a single institution using multiple identical 3.0 Tesla scanners (Achieva; Philips Medical Systems, Best, The Netherlands) with a head and neck coil. Fat-saturated axial fast spin-echo T2WIs were acquired with a repetition time/echo time (TR/TE) of 6480 ms/70 ms and section thickness of 4 mm before the injection of contrast material. All images were obtained with a 22–25-cm field-of-view. Fat-saturated axial T1-weighted high-resolution isotropic volume examination (THRIVE) images were acquired at 40 s after the administration of the contrast material and then reformatted to the coronal and sagittal planes. The parameters for the fat-saturated 3D THRIVE images were as follows: TR=4.5 ms, TE=2.2 ms, flip angle=10°, bandwidth=434 Hz/pixel, matrix=256×256, section thickness=1 mm, signal averaging=2, and acquisition time=223 s.

Tumor segmentation was performed by a radiologist (J.K., with 19 years of experience in head and neck imaging) on the 3D CE-T1WIs for the selection of the contrast-enhanced tumor using semi-automatic methods, such as signal intensity threshold, region growing, and edge detection. After image preprocessing, a total of 202 radiomic features were extracted. The details of the preprocessing and radiomic feature extraction are described in Supplementary Material 1 (only online).

Additionally, we performed visual assessment and evaluated the diagnostic performance of two head and neck radiologists in distinguishing SCC from lymphoma (S.B. and J.K., with 8 and 19 years of experience, respectively). After anonymization and data randomization, the readers were given two image sets for each patient, which included T2WIs and CE-T1WIs of the segmented tumor mask that had been used for the radiomics model. The readers recorded a final diagnosis using a 4-point scale (1=definite SCC; 2=likely SCC; 3=likely lymphoma; and 4=definite lymphoma).

All statistical analyses, including machine learning, were performed with the use of R software (version 3.5.1, R Foundation for Statistical Computing, Vienna, Austria). Values of $p<0.05$ were considered statistically significant. Two-sided statistical tests were conducted.

An elastic-net regularized generalized linear model (GLM) analysis was performed to train the machine learning classifier in this study. GLM is an extension of a linear model to vari-

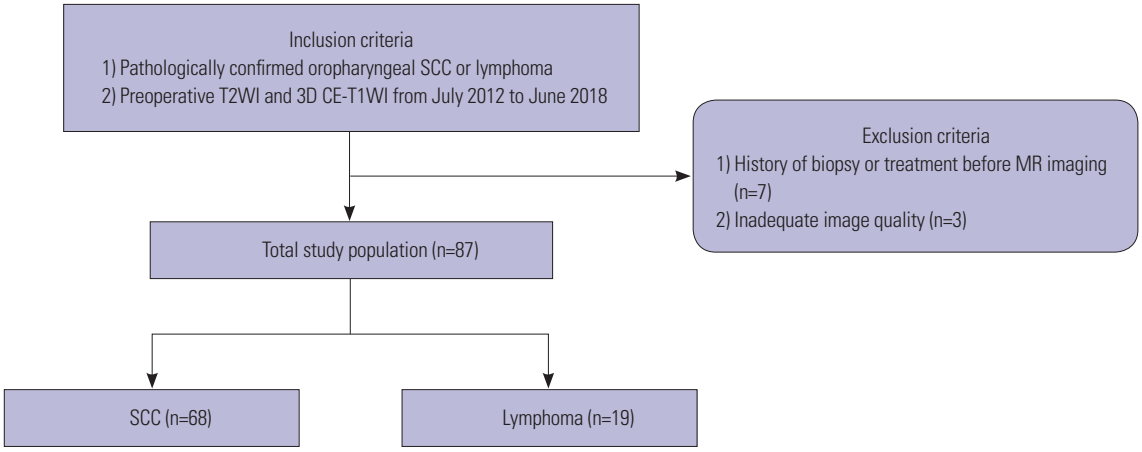

Fig. 1. Flowchart of patient enrollment. SCC, squamous cell carcinoma; T2WI, T2-weighted imaging; 3D, three-dimensional; CE-T1WI, contrast-enhanced T1-weighted imaging.

https://doi.org/10.3349/ymj.2020.61.10.895

ables that do not follow normal distribution. Elastic net regularization, a regularization method used to reduce the variance in linear regression models, has been reported to be useful when the number of predictors is considerably greater than that of the observations or when predictors are correlated as in radiomics data. Analyses were performed with nested cross-validation to avoid overfitting. The details of the nested cross-validation are presented in Supplementary Material 2 (only online). The elastic-net regularized GLM was performed using the "glmnet" package in R software. A schematic diagram of the workflow (from image processing to machine learning) is depicted in Fig. 2.

A total of 102 radiomic features were selected at least once during cross-validation, with an average of 27.6 features selected for each fold. Among these, 19 features consisting of 10 first-order features, and 9 texture features were selected at least twice. The generated heatmap of the 19 radiomic features selected for the differentiation between SCC and lymphoma yielded different distributions between two pathologies (Fig. 3). The details, including the number of times each feature was selected and the mean coefficient values of the selected features during cross-validation, are summarized in Supplementary Table 1 (only online).

The mean AUC of the radiomics classifier was 0.750 [95% confidence interval (CI), 0.613–0.887], with a sensitivity of 84.2%, a specificity of 60.3%, and an accuracy of 65.5% (Fig. 4). Human reader 1 yielded an AUC of 0.613 (95% CI, 0.467–0.759), with a sensitivity of 61.8%, a specificity of 52.6%, and an accuracy of 59.8%. Reader 2 yielded an AUC of 0.663 (95% CI, 0.531–0.795), with a sensitivity of 60.3%, a specificity of 63.2%, and an accuracy of 60.9%.

In this study, we assessed the diagnostic value of radiomic features from MR imaging in differentiating between SCC and lymphoma of the oropharynx. In our study, we used conventional MR imaging sequences, which are generally employed for head and neck MR imaging in daily clinical practice, and a radiomics-based machine learning model exhibited fair performance. In addition, we used open-source tools, such as python and R, instead of in-house software, to increase the availability and reproducibility of the radiomics-based machine learning analyses in external centers.

Although SCC and lymphoma are the most common malignancies in the oropharynx, their conventional imaging findings are sometimes indistinguishable, as reported previously.[4] In the sinonasal cavity, lymphomas are thought to yield larger tumor volume with more homogeneous signal intensity and less intratumoral necrosis than SCC.[14,15] These findings may be associated with a prolonged hypoxic microenvironment in SCC and consequent tumor necrosis.[16] In addition, several studies have discriminated between SCC and lymphoma with advanced MR imaging techniques, such as diffusion and perfusion MR imaging. Several investigators reported that lymphoma yields lower apparent diffusion coefficient values than SCC due to higher degrees of cellularity.[2-4] A dynamic perfusion study indicated that lymphomas of the oropharynx have lower K^{trans} and lower v_e values than SCC, which means that lymphoma is associated with lower vascular permeability and less extravascular extracellular space.[6] Therefore, we hypothe-

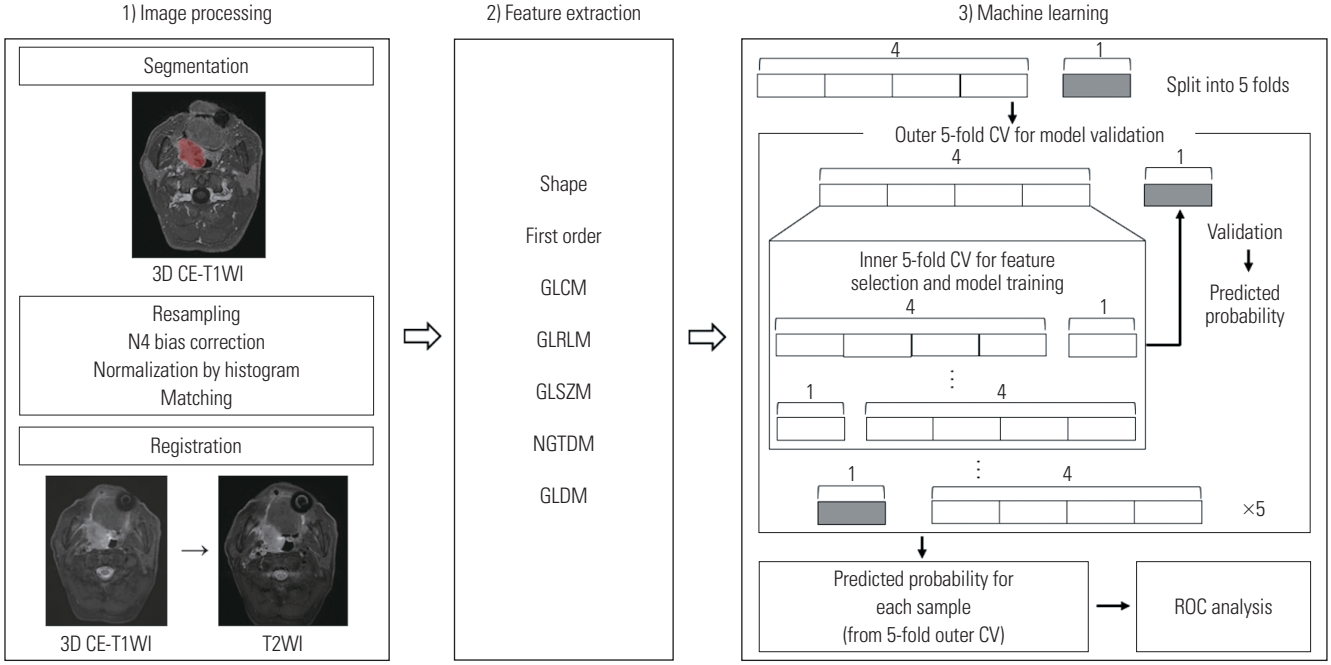

Fig. 2. Workflow of imaging processing, feature extraction, and machine learning. T2WI, T2-weighted imaging; 3D, three-dimensional; CE-T1WI, contrast-enhanced T1-weighted imaging; GLCM, gray level co-occurrence matrix; GLRLM, gray level run length matrix; GLSZM, gray level size zone matrix; NGTDM, neighboring gray tone difference matrix; GLDM, gray level dependence matrix; CV, cross-validation; ROC, receiver operating characteristic.

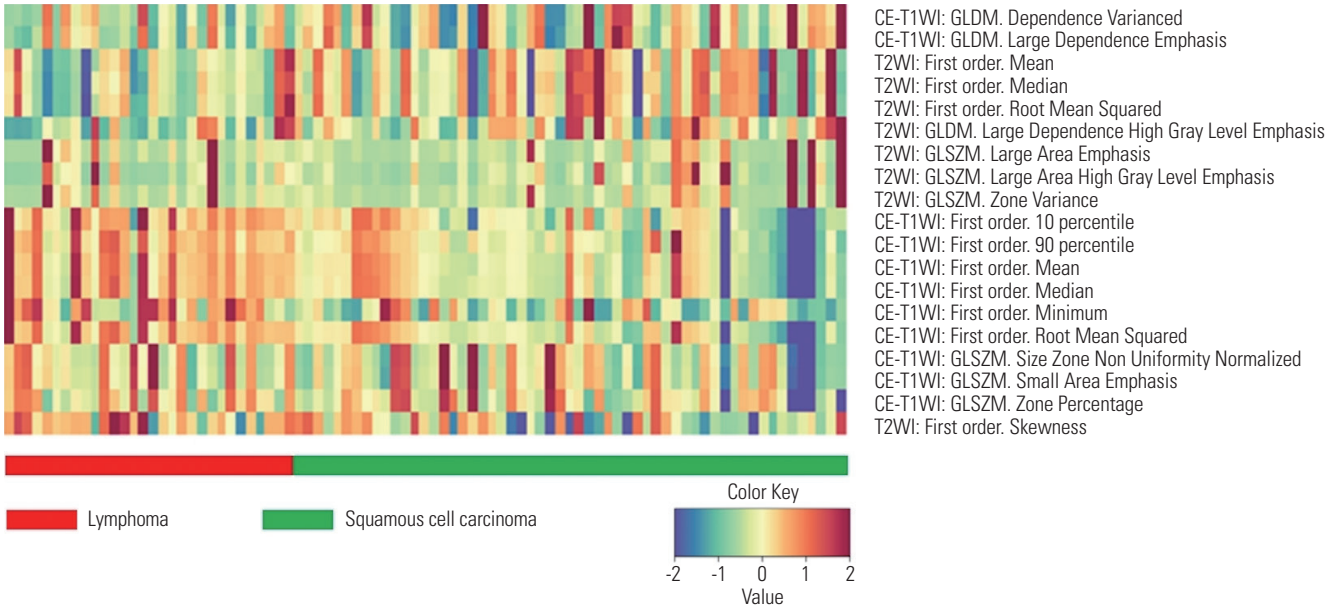

Fig. 3. Heatmap of the 19 radiomic features that were selected for the radiomic classifier at least twice during nested cross-validation to differentiate between squamous cell carcinoma and lymphoma. Each column represents individual patients, and each row represents the color-coded z-score of 19 normalized radiomic features respectively.

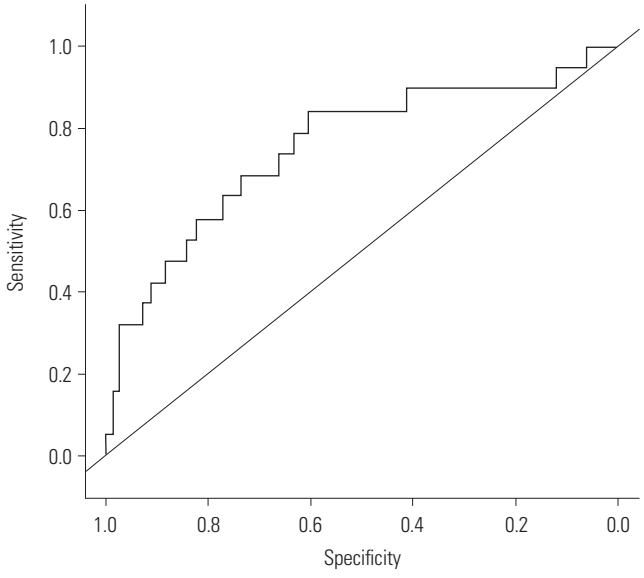

Fig. 4. Receiver operating characteristic curve of the radiomics elastic-net regularized generalized linear model analysis. The mean area under the curve for the radiomics classifier was 0.750.

sized that these microscopic differences could be reflected by quantitative radiomic features extracted from conventional MR imaging.

The 19 relevant radiomic features that were selected by the nested cross-validation in our study consisted of 10 first-order features and 9 texture features. First-order features represent voxel intensity, which is acquired from histogram-based methods,[9] and homogeneous enhancement and low T2 signal intensity in lymphomas may mainly contribute to first-order features in radiomics analyses.[17] Meanwhile, texture features describe the statistical interrelationships between voxels and reflect intratumoral heterogeneity.[9] The increased heterogeneity of the tumor signal intensity with more frequent intratumoral necrosis and hemorrhage in SCC may be the main contributor to the texture features.[14,15] However, although radiomic features have been shown to capture variable microscopic characteristics of the tumor, there are limitations in the understanding of their underlying biologic mechanism and in the identification of the value of individual features.[9,18]

In fact, in daily clinical practice, radiologic diagnosis is usually based on the integration of various information obtained from imaging studies. Radiologists obtain the information from intrinsic tumor characteristics and from other associated imaging features, such as tumor growth patterns, adjacent structural changes, and metastatic lymph node appearance. For example, adjacent bone destruction is more prominent in SCC than in lymphoma,[14] and central necrosis is one of the characteristic findings of metastatic lymph nodes from SCC.[15] Therefore, additional radiomics information obtained from the intrinsic tumor appearance would be useful to improve the diagnostic performance of radiologists.

A previous study assessed stability and prognostic performance in various combinations of feature selection and machine learning models in head and neck cancers.[19] They suggested three classifiers; Bayesian, random forest, and nearest neighbor, as the preferred methods in radiomics-based prognostic analyses, owing to their high performances and stabilities. The GLM model that was used in our study yielded high-prognostic power but low stability, while the best classifier model could vary according to the clinical outcomes, imaging modalities, and radiomic cohorts.[19,20] Large-scale comparative

https://doi.org/10.3349/ymj.2020.61.10.895

studies dealing with various feature selection and machine learning methods across independent cohorts are required to identify the optimal and most reliable estimator with which to differentiate between SCC and lymphoma of the oropharynx.

Our study has several limitations. First, the dataset was small and imbalanced with 68 SCCs and 19 lymphomas. Overfitting is an important issue in machine learning when dealing with high-dimensional features and small populations. We attempted to mitigate overfitting by nested cross-validation. Also, data imbalances can have a negative effect on the fitting of the classification algorithm because learning algorithms tend to be biased toward the majority class. We adopted a resampling technique, the synthetic minority over-sampling technique (SMOTE), to overcome the data imbalance in this study. However, this technique does not always guarantee higher performance: sometimes it even results in lower performance according to the classifier used, compared with methods with no resampling.[21-23] Second, although the model in the present study exhibited fair performance in differentiating between SCC and lymphoma of the oropharynx, it cannot obviate the need for biopsy at this point. Further study with a larger population and more elaborate algorithms could support the potential of radiomics in noninvasive diagnosis of oropharyngeal tumors. Third, advanced MR imaging techniques, such as diffusion-weighted imaging and perfusion MR imaging, were not included in this study. Although multiparametric analysis would increase diagnostic accuracy in differentiating between SCC and lymphoma, advanced techniques require additional scans and their protocols vary among institutions. Fourth, tumors were semi-automatically segmented, which is user-dependent and time-consuming. Automated tumor segmentation based on deep learning would allow further automation of the workflow, minimize user bias, and enable larger-scale studies. Last, we did not test the model in independent validation cohorts. Further studies with larger multicenter datasets, such as the Cancer Genome Atlas dataset, would be required to assess the feasibility and allow the utilization of this technology in clinical practice.

In conclusion, the radiomics-based machine learning model can be useful for differentiating SCC from lymphoma of the oropharynx, although it cannot obviate the need for invasive biopsy procedures presently. Future investigations with larger datasets and more elaborate algorithms are necessary for augmentation of the diagnostic performance of radiologists.

ACKNOWLEDGEMENTS

This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean Government (Ministry of Science and ICT) (No. 2019R1A2C1008409).

AUTHOR CONTRIBUTIONS

Conceptualization: Sohi Bae and Jinna Kim. Data curation: Sohi Bae, Beomseok Sohn, and Jinna Kim. Formal analysis: Yoon Seong Choi and Beomseok Sohn. Funding acquisition: Sung Soo Ahn, Seung-Koo Lee, Jaemoon Yang, and Jinna Kim. Investigation: Sohi Bae, Beomseok Sohn, and Jinna Kim. Methodology: Sohi Bae, Yoon Seong Choi, Beomseok Sohn, Sung Soo Ahn, and Jinna Kim. Project administration: Sung Soo Ahn, Seung-Koo Lee, Jaemoon Yang, and Jinna Kim. Resources: Sohi Bae, Beomseok Sohn, Seung-Koo Lee, and Jinna Kim. Software: Yoon Seong Choi, Beomseok Sohn, Jaemoon Yang, and Sung Soo Ahn. Supervision: Sung Soo Ahn, Seung-Koo Lee, and Jinna Kim. Visualization: Yoon Seong Choi, Beomseok Sohn, and Sung Soo Ahn. Writing—original draft: Sohi Bae, Yoon Seong Choi, and Beomseok Sohn. Writing—review & editing: Sung Soo Ahn, Seung-Koo Lee, Jaemoon Yang, and Jinna Kim. Approval of final manuscript: all authors.

ORCID iDs

Sohi Bae https://orcid.org/0000-0002-5048-4627
Yoon Seong Choi https://orcid.org/0000-0002-6034-9912
Beomseok Sohn https://orcid.org/0000-0002-6765-8056
Sung Soo Ahn https://orcid.org/0000-0002-0503-5558
Seung-Koo Lee https://orcid.org/0000-0001-5646-4072
Jaemoon Yang https://orcid.org/0000-0001-7365-0395
Jinna Kim https://orcid.org/0000-0002-9978-4356

REFERENCES

1. Harnsberger HR, Bragg DG, Osborn AG, Smoker WR, Dillon WP, Davis RK, et al. Non-Hodgkin's lymphoma of the head and neck: CT evaluation of nodal and extranodal sites. AJR Am J Roentgenol 1987;149:785-91.
2. Ichikawa Y, Sumi M, Sasaki M, Sumi T, Nakamura T. Efficacy of diffusion-weighted imaging for the differentiation between lymphomas and carcinomas of the nasopharynx and oropharynx: correlations of apparent diffusion coefficients and histologic features. AJNR Am J Neuroradiol 2012;33:761-6.
3. Maeda M, Kato H, Sakuma H, Maier SE, Takeda K. Usefulness of the apparent diffusion coefficient in line scan diffusion-weighted imaging for distinguishing between squamous cell carcinomas and malignant lymphomas of the head and neck. AJNR Am J Neuroradiol 2005;26:1186-92.
4. Sumi M, Ichikawa Y, Nakamura T. Diagnostic ability of apparent diffusion coefficients for lymphomas and carcinomas in the pharynx. Eur Radiol 2007;17:2631-7.
5. Abdel Razek AAK. Arterial spin labelling and diffusion-weighted magnetic resonance imaging in differentiation of recurrent head and neck cancer from post-radiation changes. J Laryngol Otol 2018;132:923-8.
6. Park M, Kim J, Choi YS, Lee SK, Koh YW, Kim SH, et al. Application of dynamic contrast-enhanced mri parameters for differentiating squamous cell carcinoma and malignant lymphoma of the oropharynx. AJR Am J Roentgenol 2016;206:401-7.
7. Bernstein JM, Homer JJ, West CM. Dynamic contrast-enhanced magnetic resonance imaging biomarkers in head and neck cancer: potential to guide treatment? A systematic review. Oral Oncol 2014;50:963-70.
8. Asaumi J, Yanagi Y, Konouchi H, Hisatomi M, Matsuzaki H, Kishi K. Application of dynamic contrast-enhanced MRI to differentiate

malignant lymphoma from squamous cell carcinoma in the head and neck. Oral Oncol 2004;40:579-84.
9. Gillies RJ, Kinahan PE, Hricak H. Radiomics: images are more than pictures, they are data. Radiology 2016;278:563-77.
10. Lambin P, Leijenaar RTH, Deist TM, Peerlings J, de Jong EEC, van Timmeren J, et al. Radiomics: the bridge between medical imaging and personalized medicine. Nat Rev Clin Oncol 2017;14:749-62.
11. Dang M, Lysack JT, Wu T, Matthews TW, Chandarana SP, Brockton NT, et al. MRI texture analysis predicts p53 status in head and neck squamous cell carcinoma. AJNR Am J Neuroradiol 2015;36:166-70.
12. Ramkumar S, Ranjbar S, Ning S, Lal D, Zwart CM, Wood CP, et al. MRI-based texture analysis to differentiate sinonasal squamous cell carcinoma from inverted papilloma. AJNR Am J Neuroradiol 2017;38:1019-25.
13. Ger RB, Zhou S, Elgohari B, Elhalawani H, Mackin DM, Meier JG, et al. Radiomics features of the primary tumor fail to improve prediction of overall survival in large cohorts of CT- and PET-imaged head and neck cancer patients. PLoS One 2019;14:e0222509.
14. Kim SH, Mun SJ, Kim HJ, Kim SL, Kim SD, Cho KS. Differential diagnosis of sinonasal lymphoma and squamous cell carcinoma on CT, MRI, and PET/CT. Otolaryngol Head Neck Surg 2018;159:494-500.
15. Kato H, Kanematsu M, Watanabe H, Kawaguchi S, Mizuta K, Aoki M. Differentiation of extranodal non-Hodgkins lymphoma from squamous cell carcinoma of the maxillary sinus: a multimodality imaging approach. Springerplus 2015;4:228.
16. Li JZ, Gao W, Chan JY, Ho WK, Wong TS. Hypoxia in head and neck squamous cell carcinoma. ISRN Otolaryngol 2012;2012:708974.
17. Derinkuyu BE, Boyunağa Ö, Öztunalı Ç, Tekkeşin F, Damar Ç, Alımlı AG, et al. Imaging features of Burkitt lymphoma in pediatric patients. Diagn Interv Radiol 2016;22:95-100.
18. Lambin P, Rios-Velazquez E, Leijenaar R, Carvalho S, van Stiphout RG, Granton P, et al. Radiomics: extracting more information from medical images using advanced feature analysis. Eur J Cancer 2012;48:441-6.
19. Parmar C, Grossmann P, Rietveld D, Rietbergen MM, Lambin P, Aerts HJ. Radiomic machine-learning classifiers for prognostic biomarkers of head and neck cancer. Front Oncol 2015;5:272.
20. Mackin D, Fave X, Zhang L, Fried D, Yang J, Taylor B, et al. Measuring computed tomography scanner variability of radiomics features. Invest Radiol 2015;50:757-65.
21. Xie C, Du R, Ho JW, Pang HH, Chiu KW, Lee EY, et al. Effect of machine learning re-sampling techniques for imbalanced datasets in 18F-FDG PET-based radiomics model on prognostication performance in cohorts of head and neck cancer patients. Eur J Nucl Med Mol Imaging 2020 Apr 6 [Epub]. Available at: https://doi.org/10.1007/s00259-020-04756-4.
22. Park YW, Oh J, You SC, Han K, Ahn SS, Choi YS, et al. Radiomics and machine learning may accurately predict the grade and histological subtype in meningiomas using conventional and diffusion tensor imaging. Eur Radiol 2019;29:4068-76.
23. Park YW, Choi YS, Ahn SS, Chang JH, Kim SH, Lee SK. Radiomics MRI phenotyping with machine learning to predict the grade of lower-grade gliomas: a study focused on nonenhancing tumors. Korean J Radiol 2019;20:1381-9.

https://doi.org/10.3349/ymj.2020.61.10.895