

# Prediction of squamous cell carcinoma cases from squamous cell hyperplasia in throat lesions using CT radiomics model

Osama A. Khodrog, PhD, Fengzhi Cui, MS, Nannan Xu, MD, Qinghe Han, PhD, Jianhua Liu, PhD, Tingting Gong, MD, Qinghai Yuan, PhD.

## ABSTRACT

**الأهداف:** خلصت هذه الدراسة بمجموعة من الأهداف ذات الأهمية والتي تتمثل باستخراج ميزات راديوميكس الإشعاعي من صور الأشعة المقطعية المصاحبة والغير مصاحبة للصبغة (المرحلة الوريديّة والمرحلة الشريانيّة) للحنجرة، فضلاً عن عمل التحليلات الكمية باستخدام أنواع الصور الثلاثة المصاحبة والغير مصاحبة للصبغة من أجل التفريق بين سرطان الخلايا الحرشفية وتضخم الخلايا الحرشفية في الحنجرة، إضافة إلى ذلك، تم تقييم دقة نموذج راديوميكس الإشعاعي المقترح بناء على أنواع الصور المقطعية الثلاثة التي تم بنائها في هذه الدراسة.

**المنهجية:** أما فيما يتعلق بالطرق والأدوات: فقد تم اختيار مئة مريض بأثر رجعي من المصابين الذين يعانون من سرطان الحنجرة، تم فصل المرضى الي مجموعتين هما مجموعة التدريب و تتكون من 70 حالة و مجموعة التحقق تتكون من 30 حالة، تم استخدام برنامج الذكاء الاصطناعي (The Artificial Intelligence Kit software (A. K. software) لاستخراج ميزات راديوميكس الإشعاعي من صور الأشعة المقطعية. تمت معالجة وفترة جميع الميزات باستخدام طرق (لاسو و mRMR، و LASSO. وبذلك، تم تقييم نموذج راديوميكس الإشعاعي بناء على قيم الدقة، الخصوصية، الحساسية و المساحة تحت المنحنى باستخدام برنامج R-studio.

**النتائج:** نتائج الدراسة كانت مهمة إذ أن قيم دقة التشخيص، والخصوصية، والقيمة التنبؤية الإيجابية، والقيمة التنبؤية السلبية، والمساحة تحت المنحنى التي تم الحصول عليها لمجموعة التدريب 0.91، 0.9، 0.93، 0.9 و 0.96 على التوالي للأشعة المقطعية الشريانية (CTA). وبالنسبة للأشعة المقطعية غير المصاحبة للصبغة (CTN) فقد كانت نتائجها 0.93، 0.93، 0.95، 0.90 و 0.96 على التوالي. أما الأشعة المقطعية الوريديّة (CTV) كانت نتائجها 0.92، 0.87، 0.91، 0.96 و 0.96 على التوالي. بالتالي تم تأكيد هذه القيم فيما بعد من خلال مجموعة التحقق من الصحة التي تم استخدامها للتأكيد على النتائج التي تم الحصول عليها فيما قبل.

**الخلاصة:** فقد جاءت هذه الدراسة بخلاصة ذات أهمية مفادها إن نموذج التنبؤ القائم على راديوميكس الإشعاعي المقترح في هذه الدراسة استطاع التمييز بنجاح بين سرطان الخلايا الحرشفية من تضخمها داخل الحنجرة باستخدام تقنيات التصوير المقطعي التي تم تأسيسها في هذه الدراسة، مما يؤدي إلى تطوير التشخيص قبل الجراحة على أساس المؤشرات الحيوية المحددة والأنماط الظاهرية للسرطان.

**Objectives:** To differentiate squamous cell hyperplasia (SCH) (benign) from squamous cell carcinoma (SCC) malignant) using textural features extracted from CT images and thereby, facilitate the preoperative medical diagnosis and treatment of throat cancers without the need for sample biopsies.

**Methods:** In total, 100 throat cancer patients were selected for this retrospective study. The cases were

collected from the Second Hospital of Jilin University, Changchun, China, from June 2017 to January 2019. The patients were separated into a training and validation cohort consisting of 70 and 30 cases, respectively. The Artificial Intelligence Kit software (A.K. software) was used to extract the radiomics features from the CT images. These features were further processed using the minimum redundancy maximum relevance (mRMR) and least absolute shrinkage and selection operator (LASSO) methods to obtain a subset of optimal features. The radiomics model was validated based on area-under-the-curve (AUC) values, accuracy, specificity, and sensitivity using the R-studio software.

**Results:** The diagnostic accuracy, specificity, PPV, NPV, and AUC values obtained for the training cohort was 0.91, 0.9, 0.93, 0.9, and 0.96 CT angiography (CTA), 0.93, 0.93, 0.95, 0.90, and 0.96 computed tomography normal (CTN), and 0.92, 0.87, 0.91, 0.96, and 0.96 CT venogram (CTV). These values were subsequently confirmed in the validation cohort.

**Conclusion:** The radiomics-based prediction model proposed in this study successfully differentiated between SCH and SCC throat cancers using CT imaging, thereby facilitating the development of accurate preoperative diagnosis based on specific biomarkers and cancer phenotypes.

**Keywords:** radiomics, tomography x-ray computed, nomograms, neck, machine learning

*Saudi Med J* 2021; Vol. 42 (3): 284-292  
doi: 10.15537/smj.2021.42.3.20200617

From the Department of Radiology (Khodrog, Cui, Xu, Han, Liu, Gong, Yuan), the Second Hospital of Jilin University, Changchun, China and from the Department of Medical Imaging (Khodrog), Faculty of Applied Medical Health, Palestine Ahliya University, Bethlehem, Palestine.

Received 2nd October 2020. Accepted 1st February 2021.

Address correspondence and reprint request to: Dr. Qinghai Yuan, Department of Radiology, Norman Bethune College of Medicine, The Second Hospital of Jilin University, Changchun, China. E-mail: yuanqinghai123@sina.com  
ORCID ID: <http://orcid.org/0000-0002-5337-5354>

Head and neck squamous cell carcinomas (HNSCC) represent the fifth most common cancers in the world as compared to other types of cancers.<sup>1,2</sup> The HNSCC group of cancers includes oral, nasopharyngeal, and throat cancers,<sup>3</sup> in which throat cancer is the second most common cancer affecting the respiratory tract with a survival rate of 64.2%. In general, cancer therapy includes surgery, radiotherapy, chemotherapy, or biological agents that have been selected according to the patient's case. The risk factors for throat cancer infection increases with alcohol consumption and cigarette smoking, in addition to other factors such as genetics and nutrition.<sup>4</sup> For throat cancer diagnosis, a CT scan is usually recommended by the radiologist as it is the fastest method to obtain a result. The use of thin-slice helical CT with 2-dimensional (2D) reconstruction is commonly employed for the diagnosis. During this scanning process, the appearance of soft tissue and bony details such as calcification, abnormal changes in the throat, and site of the spinal canal can be clearly demonstrated.<sup>5</sup> The differentiation between cystic or necrotic tissues is supported by CT angiography (CTA) and CT venogram (CTV) scans, whereby the tumor regions appear brighter due to their heterogeneous enhancement pattern as opposed to the normal regions.<sup>6</sup> Nevertheless, the radiologist faces a challenge in the diagnosis of throat cancers especially in the early-stages due to non-specific symptoms such as hemorrhage, calcification, and presence of heterogeneous masses. The tumor histology is usually determined using biopsy, although this method may produce errors in sampling or region detection and subsequently result in complications.<sup>7</sup> Therefore, using CT imaging, selected features from the clinical images of the tumor can be extracted to discriminate intra-tumor heterogeneity. The differentiation between the intratumor heterogeneity has a significant effect on the prediction, diagnosis, prognosis, response, and staging of cancer.<sup>8</sup> Radiomics is a powerful technique

employed to analyze image textures. It is a new method in translational research that has been applied for the discrimination of cancer phenotypes using quantitative features derived from the imaging data. For instance, a high number of textural features such as measurements from the mean and standard deviation can be extracted from the imaging data and converted into predictive mathematical models<sup>9</sup> that may facilitate clinical decisions. Radiomics is a useful method to differentiate between benign and malignant lesions in head and neck cancers, and early metastatic lesions as well as provide prognosis and predictive factors for the tumor, thereby leading to an effective treatment plan for patients.

Usually, throat cancers are discovered in the later stages since it is difficult to diagnose this cancer in the early stages, thus decreasing the rate of survival. Therefore, radiomics is applied to the medical imaging analyses on existing tumors or newly diagnosed tumors of patients as it may provide new information to physicians. For instance, the expression of genomic and proteinuria systems based on the features of the radiomics model for these images can be extracted; thus, allowing physicians to obtain phenotypic gene-protein patterns which may include prognostic information.<sup>10,11</sup> Therefore, this additional information is beneficial and provides details that may not be visible to the naked eye. This study aims to differentiate squamous cell hyperplasia (SCH) (benign) from squamous cell carcinoma (SCC) (malignant) using textural features extracted from CT images and thereby, facilitate the preoperative medical diagnosis and treatment of throat cancers without the need for sample biopsies.

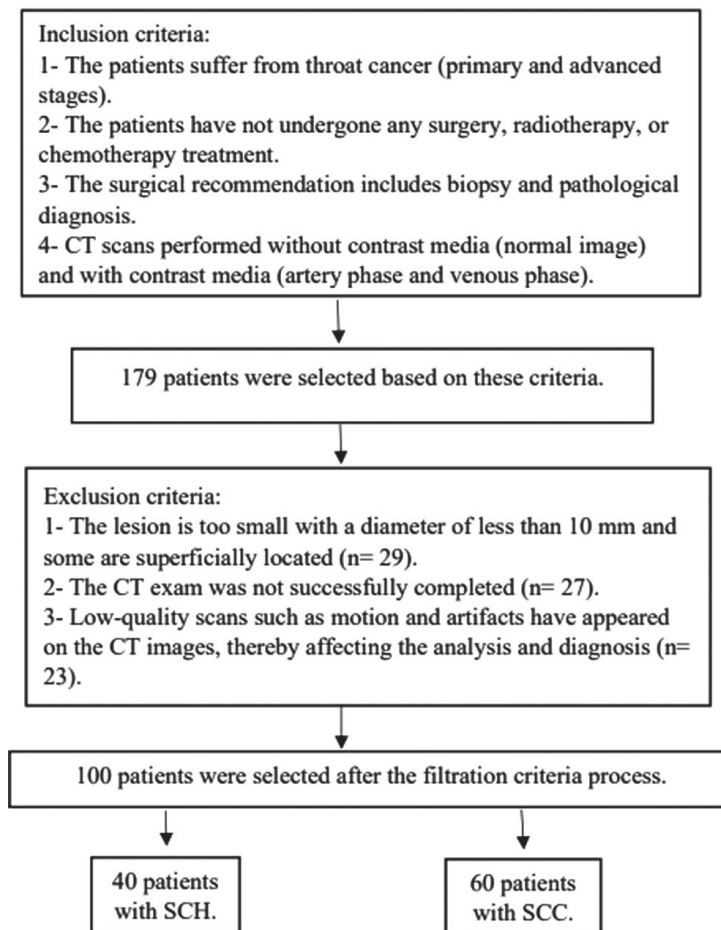
**Methods.** This retrospective throat cancer diagnostic study was approved by the Second Hospital Ethical Committee of Jilin University, Changchun, China. The informed consent was waived off because of the nature of the study, which was retrospective, with preserving patients' data and not disclosing and handled with strictly confidential. This study has been conducting based on the Helsinki Declaration. The medical imaging technology used in this study was based on the picture archiving and communication system (PACS) obtained from the Second Jilin University Hospital, Changchun, China. The images were reviewed from June 2017 to January 2019 to select patients who were pathologically diagnosed with throat cancer. This study focused on patients that underwent CT scanning, in which the medical history of patients such as gender, smoking, alcohol, tumor staging, and histopathology of the surgical specimen was available in the database. In total, 100 patients with throat cancer were selected

**Disclosure.** Authors have no conflict of interests, and the work was not supported or funded by any drug company. This study was supported by the Technology Development Plan for Jilin Province, China (No. 2018010119JC), the Youth Training Foundation by the Health and Family Planning Commission, Jilin Province, China (No. 2017Q012), and the Health Special Project of Jilin Province Department of Finance, Jilin Province, China (No. 3D518V283429).

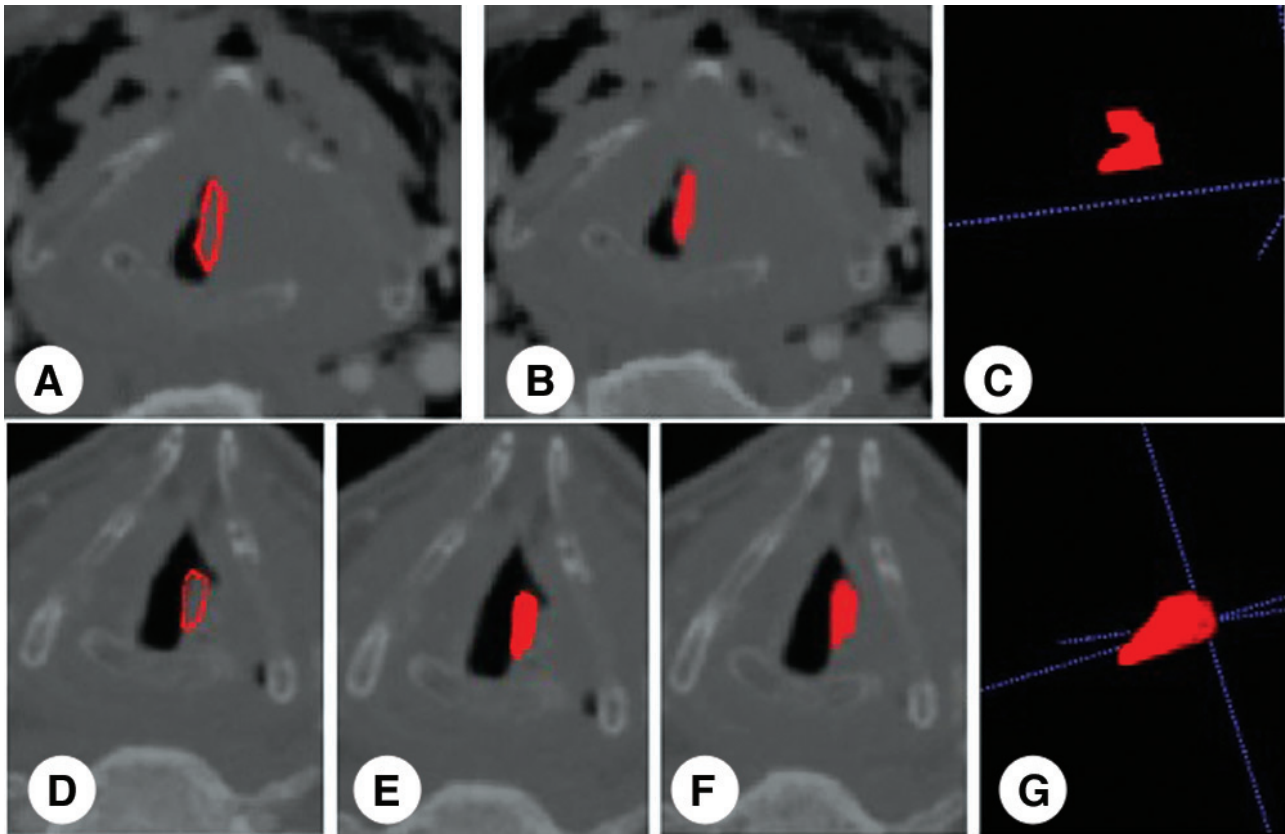
for the training and validation groups, in which 60 patients were diagnosed with SCC and 40 patients were diagnosed with SCH. Based on their gender and age group classification, 73 males and 27 females, ranging from 44 to 85 years, were selected as participants and the Kolmogorov-Smirnov test was applied to determine the normal distribution of the data. Information regarding the lifestyle choices of the patients was also recorded and the patient profile consisted of smokers (n=74), non-smokers (n=14), and those who quit smoking (n=12). In addition, the participants also comprised alcoholic-drinkers (n=73), non-drinkers (n=20) and those who stopped drinking (n=7). This study established exclusion and inclusion criteria and followed with high strictness and transparency to know these criteria. **Figure 1** depicts the classification of the study participants and clarifies the mechanism followed to exclude and include patients.

### *Computerized tomography (CT) scan parameters.*

The CT scans of the throat were performed using a 256-Slice scanner (Philips iCT, Netherland), with the tube voltage and tube current set at 120 kVp and 200 mAs, respectively. The scan parameters were established as follows: scanning layer thickness of 1 mm in diameter, scanning layer spacing of 1 mm, pitch ratio of 0.342, matrix array dimension of 256 x 256, and scan time of 1-3 seconds. A preoperative throat scan was first obtained using CTN imaging that was performed without contrast media. The second CT scan was accomplished for CTA 35s and CTV 65s using contrast media as well as a non-ionic iodinated contrast medium for intravenous injection (dosage concentration of 1.5 mL/kg and injection rate of 3.5 mL/s) (Ultravist, 100 Bayer Blvd, Whippany, NJ, United States). The reconstruction was performed for all the throat scans



**Figure 1** - Flowchart representing the patient selection criteria process. SCH: squamous cell hyperplasia, SCC: squamous cell carcinoma (SCC)



**Figure 2** - The segmentation process by ITK-SNAP software for squamous cell carcinoma showing A-C) malignant and D-G) the squamous cell hyperplasia benign.

according to the established criteria. The CT images were exported to DICOM format for feature extraction.

**Setup experiments.** The CT images of SCH and SCC were segmented manually based on the region of interest (ROI) by a radiologist with 3 years of experience. Subsequently, the radiomics features were extracted from these segmented CT images. The ROI was reviewed by an expert radiologist with more than 20 years of experience. The ROI was drawn around the tumor outline walls for all the images. The borders of the tumor were excluded from this study as the edges of the tumor can sometimes merge with normal tissues and subsequently produce weak results. The ROI's artificial outline for the CT images was performed using ITK-SNAP 3.8.0 ([www.itksnap.org](http://www.itksnap.org))<sup>12,13</sup> (Figure 2).

The radiomics features obtained using the A.K. software (GE Healthcare, China) were automatically calculated and divided into volumetric, texture, histogram, and wavelet features.<sup>14</sup> The radiomics data provided 396 features, although most of them were not useful for clinical applications. For the selection of

useful features with potential clinical application, least absolute shrinkage and selection operator (LASSO) and minimum redundancy maximum relevance (mRMR) methods were employed to eliminate redundant and irrelevant features before proceeding with the construction of the logistic regression model. The mRMR method was used first, and followed by the LASSO method which was used as a confirmation method. These features were associated with one or more useful outcomes under the logistic regression model. These features were subsequently divided into 6 groups consisting of a histogram, gray level size zone matrix, form factor matrix, gray-level co-occurrence matrices, and run-length matrix for data analysis.

For the selection process using the mRMR method, 30 features were retained in this process and the irrelevant features were eliminated. Next, the LASSO method was performed to extract the most effective subset of features for the construction of the final logistic regression model. In the LASSO model, a 10-fold cross-validation analysis was used to create the

optimal penalization coefficient lambda ( $\lambda$ ). The  $\lambda$  value that gave the lowest average binomial deviance was used in the feature selection. The vertical lines accompanied by red points were drawn for the optimal values using the lowest criteria. Lastly, a coefficient profile plot was produced against the log ( $\lambda$ ) sequence based on the coefficient LASSO profiles of 1188 features.

Statistical analysis. The statistical analyses for the extracted features obtained from the CTA, computed tomography normal (CTN), and CTV images were performed using the open-source R-Studio Server software. The values for the multivariate logistic model were considered significant at  $p < 0.05$ . The receiver-operating-characteristic (ROC) and area-under-the-curve (AUC) analyses were used to evaluate these models. The logistic regression model and ROC curve were analyzed using the R software package to discriminate SCH from SCC lesions from the training cohort and validation cohort. The optimal threshold value calculated for the training cohort was applied to the validation cohort to differentiate the SCH lesions from SCC lesions. Therefore, the optimal threshold was obtained by maximizing the Youden index. All the statistical analyses were performed using R-Studio Server (Version 1.1.463, GE Healthcare, China).

**Results.** The data in this study was obtained from 100 patients who suffered from throat cancer. Based on their histology and pathology results, 60 patients suffered from SCC and the remaining 40 patients suffered from SCH. These patients were divided into the training group (70%) and test group (30%). Additionally, no statistical differences were observed between the age ( $p=0.12$ ), gender ( $p=0.50$ ), smoker status ( $p=0.90$ ), and alcohol consumption ( $p=0.90$ ) for both groups.

**Construction of a radiomics model and feature extraction from CT images.** In total, 1188 features were extracted from the CTA, CTN, and CTV

images, whereby 396 features were extracted from each segmented ROI image. Once the number of features were stabilized, the most effective subset of the extracted features was determined to evaluate the corresponding coefficients. A total of 12 features were selected from the extracted features for CTA, 11 for CTN, and 12 for CTV; thereby, indicating the construction of a promising radiomics-based prediction model for throat cancer.

**The diagnostic capability of the CT radiomics model for throat cancer.** In this study, CTA, CTN, and CTV parameters were used for the CT images to measure the diagnostic abilities of these parameters based on the receiver operating characteristic (ROC) curve values. The dataset obtained using these parameters were trained separately using the logistic regression model. The radiomics model constructed based on the CT images resulted in the AUC value of 0.96 (95% CI: 0.92-1) for all 3 parameters (CTA, CTN, and CTV) in the training cohort.

The logistic regression model was then tested using the validation cohort for all 3 parameters. The radiomics model based on the CT images resulted in the AUC value of 0.94 (95% CI: 0.82-1) for CTA and CTV, while a higher AUC value of 0.96 (95% CI: 0.87-1) was obtained for the CTN parameter.

The predictive performance of the different CT images and comparison of radiomics features extracted from these images such as accuracy, specificity, sensitivity, positive predictive value (PPV), and negative predictive values (NPV) for the training cohort and validation cohort are shown in **Table 1**.

Based on the results obtained, the CTN radiomics model provides a promising diagnostic ability for throat cancer as compared to the other parameters of the CT images. Specifically, the logistic regression model produced significantly higher results for AUC (0.96), accuracy (0.93), sensitivity (0.93), and specificity

**Table 1** - The performance of the CT parameters based on the radiomics model.

CT parameters	Training data					Testing data				
	Acc.	Sen.	Spe.	PPV.	NPV.	Acc.	Sen.	Spe.	PPV.	NPV.
CTA	0.91	0.92	0.9	0.93	0.9	0.78	0.93	0.0	0.83	0.0
CTN	0.93	0.93	0.933	0.95	0.90	0.84	0.94	0.33	0.88	0.50
CTV	0.92	0.97	0.87	0.91	0.96	0.78	0.75	1	1	0.43

CTA: computed tomography artery, CTV: computed tomography venous, CTN: computed tomography normal, Acc: accuracy, Sen: sensitivity, Spe: specificity, PPV: positive predictive value, NPV: negative predictive value

(0.93) in the training cohort, and these results were subsequently confirmed in the validation cohort with values of 0.96, 0.84, 0.94, and 0.33, respectively. However, the CTN radiomics model yielded better results due to the absence of contrast media, while the contrast enhancement displayed in CTA and CTV introduced heterogeneity in the radiomics features and decreased their predictive ability.

The mean AUC values for all 3 parameters used in this study are represented in **Figure 3**. The mean AUC value for CTN was higher compared to the values obtained for CTA and CTV imaging. A statistically significant difference was also observed between the values for all 3 parameters ( $p < 0.0001$ ).

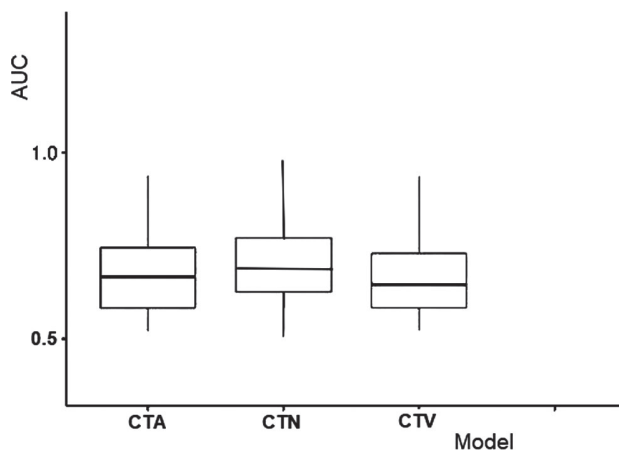
The predictive performance of the radiomics model proposed in this study was higher as compared to the clinical model. **Figure 3** depicts the nomogram for the decision curve based on the evaluation of the clinical application of the imaging data. The net benefit in the y-axis of the nomogram is dependent on the difference in the expected benefit and the expected harm for different threshold probabilities in the validation cohort based on the introduced model [Net benefit = true positive rate - (false positive rate × weighting factor), weighting factor = threshold probability/(1-threshold probability)].

**The analysis of extracted features.** Based on the feature selection by the mRMR and LASSO methods, only the most significant features identified in the logistic regression model for each parameter were used in this study. In total, 12 features (Inverse Difference Moment,

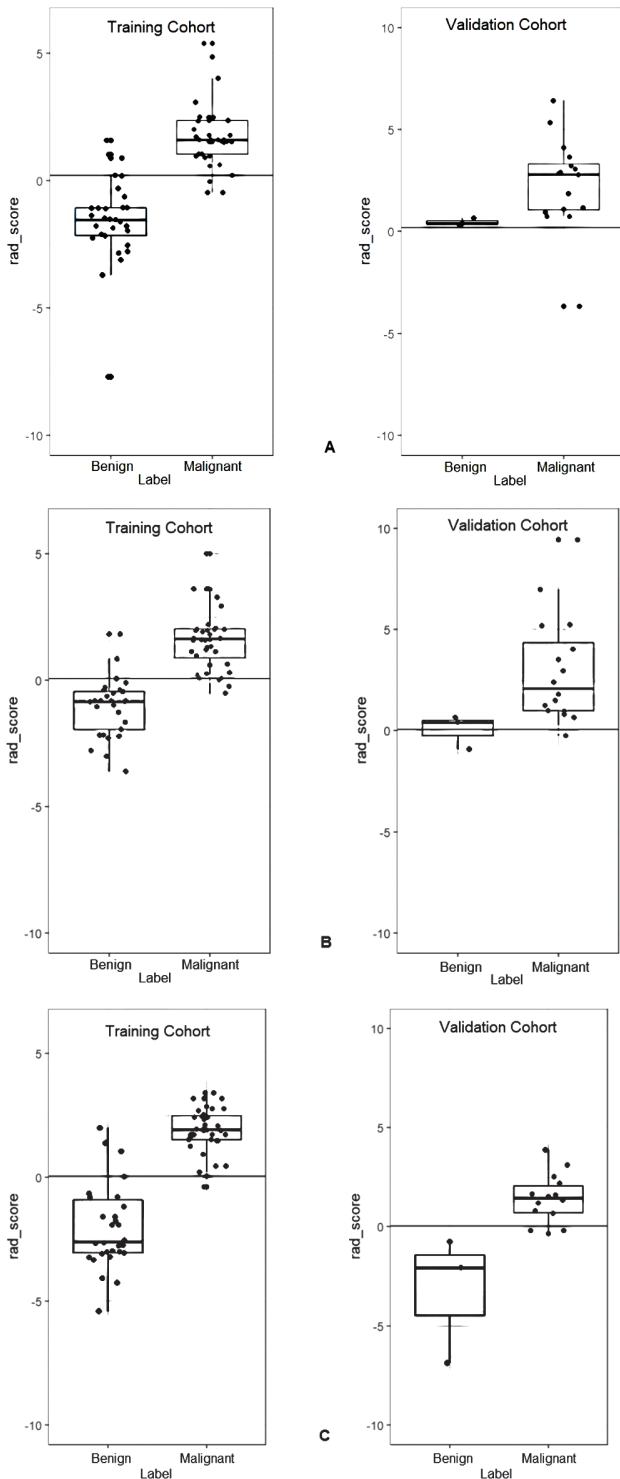
Cluster Shade, GLCM Energy, Haralick Correlation, Volume, Short Run Emphasis, Long Run High Grey Level Emphasis, Correlation, Haralick Correlation) were selected for CTA (**Figure 4A**), while 11 features (Cluster Prominence, Cluster Shade, GLCM Energy, Haralick Correlation, Correlation, Short Run Low Grey Level Emphasis, Inverse Difference Moment), and 12 features (Cluster Shade, GLCM Energy, Correlation, Inertia, Surface Area, Run Length Nonuniformity, GLCM Entropy) were selected for CTN (**Figures 4B**) and CTV (**Figure 4C**), respectively. The features selected using the LASSO model parameters varied significantly between SCH lesion and SCC. The statistically significant values for the training cohort and validation cohort are listed in **Table 2**.

**Discussion.** In this study, 3 parameters (CTA, CTN, and CTV) were used for the CT images based on the logistic regression model as compared to a previous study that used only one parameter.<sup>15</sup> The results obtained in this study yielded statistically significant differences between the SCH and SCC lesions based on the 3 parameters used as compared to previous studies.<sup>16,17</sup> In these studies, the authors proposed the use of computer-aided diagnosis (CAD) for the prediction of patients with throat cancer, thereby increasing the diagnostic performance for early-stage cancers<sup>18</sup> as well as minimizing errors using a classification system. In this study, values of 0.78 were obtained for CTA, 0.84 for CTN, and 0.78 for CTV. Data from the CTN imaging resulted in a higher accuracy, specificity, and sensitivity in the training cohort, in which the data was subsequently confirmed in the validation cohort based on the established parameters. Thus, these results indicate that the diagnostic evaluation established in this study was relatively stable. Additionally, the discrimination of SCH lesions from SCC throat cancers using CT images was supported by the development of a radiomics score. The radiomics score consisted of a nomogram that performed well in differentiating between benign lesions and malignancy. The prediction based on the calibration curve for throat cancer and the decision curve analysis (DCA), which included a nomogram, were shown to be useful for clinical application.<sup>19</sup>

Despite the importance of CTA and CTV in clinical cases, their predictive ability in the radiomics model was lower due to the use of contrast enhancement that resulted in heterogeneity for the radiomics features. No enhancement is required for CTN images, thus resulting in its higher predictive ability as compared to other parameters when used in the radiomics model.



**Figure 3** - Box plot of area-under-the-curve (AUC) values for throat cancer illustrating the higher values of computed tomography normal (CTN) parameters as compared to computed tomography artery (CTA) and computed tomography venous (CTV).



**Figure 4** - The least absolute shrinkage and selection operator (LASSO) features after univariate logistic regression and removing redundancies, for each parameter was used in our study. **A)** The right and left image of compared to computed tomography artery (CTA), **B)** the right and left image of computed tomography normal (CTN), **C)** the right and left image of computed tomography venous (CTV).

**Table 2** - The statistically significant for least absolute shrinkage and selection operator (LASSO) features selected of throat cancer for each parameter used in this study.

Model	Training cohort (p-value)	Validation cohort (p-value)
CTA	<0.000*	0.014*
CTN	<0.000*	<0.008*
CTV	<0.000*	<0.002*

\*p-value is determined by the Spearman's rank correlation test and  $p < 0.05$  is statistically significant. CTA: computed tomography artery, CTV: computed tomography venous, CTN: computed tomography normal

Radiomics is an emerging field of research based on computer-aided technology, in which medical images are converted into a sequence of data using computer algorithms.<sup>20,21</sup> The microscopic properties of the tumor biomarkers and biological phenotypes are captured by the image and the features of the tumor are extracted using radiomics software to identify the heterogeneity of the tumor biomarkers.<sup>22,23</sup> The relationship between the extracted features and biological phenotype is quite complex,<sup>24</sup> mainly due to the extraction of hundreds of radiomics features for these tumor biomarkers. Therefore, the utility of these features is evaluated based on the parameters established for the data.<sup>25</sup> The radiomics score was developed in this study to determine the prognosis of SCH lesions from SCC lesions. The differentiation, calibration, and clinical utility of this score was applied using a training cohort and subsequently, a validation cohort to evaluate the performance of this model. The radiomics model achieved a good result as compared to the conventional model.

The radiomics features derived from the CT images were quantitatively evaluated in this study to discriminate SCH lesions from SCC throat cancer lesions. Additionally, a radiomics-based prediction model was developed using the logistic regression model. The CTN scan significantly discriminated SCH lesions from SCC throat cancers, as indicated by the AUC values of 0.96 obtained in the training cohort and 0.96 in the validation cohort. The results observed in this study were consistent with previous studies indicating that normal CT images exhibited a higher prediction performance as compared to enhanced CT images.<sup>26-28</sup> The main reason for this observation is attributed to the clearer identification of the phenotypic heterogeneity of tumors using the normal CT scan as opposed to the reduced visibility of enhanced CT images due to the

existing intra-tumoral contrast material. Specifically, the ROI contains the entire lesion, as opposed to previous studies that have only  $\geq 1$  cm.<sup>29</sup> The patterns and distributions such as cystic and calcification are crucial for the differentiation between benign and malignant tumors.<sup>31</sup>

This study also offers an advantage that has not been previously described in any study. Two methods were used to select the features required for the 3 CT images, as opposed to the use of only a single method LASSO.<sup>18,29</sup> For instance, the mRMR method was performed to eliminate redundant and irrelevant features. Subsequently, the LASSO method was used to select the optimized subset.<sup>32</sup> The features were selected with high accuracy and were able to significantly discriminate SCH from SCC lesions, as evidenced by the results in this study.

**Study limitations.** The number of sample cases included in this retrospective study was insufficient as only 100 cases of SCC and SCH were analyzed. Although the Digital Imaging and Communications in Medicine (DICOM) data obtained by the single-center, single CT can ensure data consistency, the ability to performing a comprehensive evaluation of the model is lacking. Finally, the manual sketching of ROI by radiologists can lead to discrepancies. Therefore, it is envisaged that future studies should incorporate a larger sample size using multi-center radiomics data.

In conclusion, the radiomics features offer a rapid topography and biomarker assessment of cancer. Using this approach, the use of medical imaging, computer software, and machine learning are combined to obtain a clearer vision of cancer. The improvement of medical imaging data, together with the development of a new computational model, support the diagnosis and treatment of cancers with high accuracy. Cancer research is essential for the discovery of tumor heterogeneity. Hence, it is necessary to develop a computational model that optimizes the significant radiomics features for its application in clinical cases.

This study has yielded promising results for the differentiation of SCH lesions from SCC throat cancers based on the 3 parameters established for CT imaging. Moreover, the CTN data imaging was shown to have more efficiency as compared to the other parameters identified for CTA and CTV. It is envisaged that the findings of this study can facilitate the early diagnosis of throat cancer without the need for biopsies, thereby leading to higher survival rates and faster administration of medical treatments.

**Acknowledgment.** This work was supported by the Technology Development Plan for Jilin Province, China (No. 2018010119JC), the Youth Training Foundation by the Health and Family Planning Commission of Jilin Province, China (No. 2017Q012), and Health Special Project of Jilin Province Department of Finance, China (No. 3D518V283429). We would like to acknowledge Proofreading by A UK PhD company ([www.proofreadingbyphd.com](http://www.proofreadingbyphd.com)) for a marvelous English language editing service.

## References

1. Wagner S, Wittekindt C, Sharma SJ, Wuerdemann N, Jüttner T, Reuschenbach M, et al. Human papillomavirus association is the most important predictor for surgically treated patients with oropharyngeal cancer. *Br J Cancer* 2017; 116: 1604-1611.
2. Seok J, Ryu CH, Ryu J, Kim JH, Lee SJ, Park WS, et al. Prognostic implication of SOX2 expression associated with p16 in oropharyngeal cancer: A study of consecutive tissue microarrays and TCGA. *Biology (Basel)* 2020; 9: 387.
3. Diamant A, Chatterjee A, Vallières M, Shenouda G, Seuntjens J. Deep learning in head & neck cancer outcome prediction. *Sci Rep* 2019; 9: 2764.
4. Gong H, Shi Y, Xiao X, Cao P, Wu C, Tao L, et al. Alterations of microbiota structure in the larynx relevant to laryngeal carcinoma. *Sci Rep* 2017; 7: 5507.
5. Yin P, Mao N, Zhao C, Wu J, Sun C, Chen L, et al. Comparison of radiomics machine-learning classifiers and feature selection for differentiation of sacral chordoma and sacral giant cell tumour based on 3D computed tomography features. *Eur Radiol* 2019; 29: 1841-1847.
6. Avanzo M, Stancanello J, El Naqa, I. Beyond imaging: the promise of radiomics. *Physica Medica* 2017; 38: 122-139.
7. Zhou Y, Hu B, Wu Z, Cheng H, Dai M, Zhang B. The clinical outcomes for chordomas in the cranial base and spine: A single center experience. *Medicine* 2019; 98: e15980.
8. Sala E, Mema E, Himoto Y, Veeraraghavan H, Brenton JD, Snyder A, et al. Unravelling tumour heterogeneity using next-generation imaging: radiomics, radiogenomics, and habitat imaging. *Clin Radiol* 2017; 72: 3-10.
9. Constanzo J, Wei L, Tseng HH, El Naqa I. Radiomics in precision medicine for lung cancer. *Transl Lung Cancer Res* 2017; 6: 635-647.
10. Limkin EJ, Sun R, Derclé L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017; 28: 1191-1206.
11. Siddiqua U I. Radiomics. *KYAMC Journal* 2019; 10: 126-127.
12. Nolden M, Zelzer S, Seitel A, Wald D, Müller M, Franz AM, et al. The medical imaging interaction toolkit: challenges and advances : 10 years of open-source development. *Int J Comput Assist Radiol Surg* 2013; 8: 607-620.
13. Shen D, Wu G, Suk HI. Deep learning in medical image analysis. *Annu Rev Biomed Eng* 2017; 19: 221-248.
14. Fan L, Fang M, Li Z, Tu W, Wang S, Chen W, et al. Radiomics signature: a biomarker for the preoperative discrimination of lung invasive adenocarcinoma manifesting as a ground-glass nodule. *Eur Radiol* 2019; 29: 889-897.
15. Düzgün F, Tarhan S, Ovalı GY, Eskiizmir G, Pabuşçu Y. Is computed tomography perfusion a useful method for distinguishing between benign and malignant neck masses? *Ear Nose Throat J* 2017; 96: E1-E5.



16. Mao L, Chen H, Liang M, Li K, Gao J, Qin P, et al. Quantitative radiomic model for predicting malignancy of small solid pulmonary nodules detected by low-dose CT screening. *Quant Imaging Med Surg* 2019; 9: 263-272.
17. Haider SP, Burtneess B, Yarbrough WG, Payabvash S. Applications of radiomics in precision diagnosis, prognostication and treatment planning of head and neck squamous cell carcinomas. *Cancers Head Neck* 2020; 5: 6.
18. Wu L, Wang C, Tan X, Cheng Z, Zhao K, Yan L, et al. Radiomics approach for preoperative identification of stages I-II and III-IV of esophageal cancer. *Chin J Cancer Res* 2018; 30: 396-405.
19. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, et al. Reporting and interpreting decision curve analysis: A guide for investigators. *Eur Urol* 2018; 74: 796-804.
20. Gillies RJ, Kinahan PE, Hricak H. Radiomics: Images Are More than Pictures, They Are Data. *Radiology* 2016; 278: 563-577.
21. Limkin EJ, Sun R, Derclé L, Zacharaki EI, Robert C, Reuzé S, et al. Promises and challenges for the implementation of computational medical imaging (radiomics) in oncology. *Ann Oncol* 2017; 28: 1191-1206.
22. Guo Y, Hu Y, Qiao M, Wang Y, Yu J, Li J, et al. Radiomics analysis on ultrasound for prediction of biologic behavior in breast invasive ductal carcinoma. *Clin Breast Cancer* 2018; 18: e335-e344.
23. Grossmann P, Stringfield O, El-Hachem N, Bui MM, Rios Velazquez E, Parmar C, et al. Defining the biological basis of radiomic phenotypes in lung cancer. *Elife* 2017; 6: e23421.
24. Xu J, Yang P, Xue S, Sharma B, Sanchez-Martin M, Wang F, et al. Translating cancer genomics into precision medicine with artificial intelligence: applications, challenges and future perspectives. *Hum Genet* 2019; 138: 109-124.
25. 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: 749-762.
26. Orooji M, Alilou M, Rakshit S, Beig N, Khorrami MH, Rajiah P et al. Combination of computer extracted shape and texture features enables discrimination of granulomas from adenocarcinoma on chest computed tomography. *J Med Imaging (Bellingham)* 2018; 5: 024501.
27. Dennie C, Thornhill R, Sethi-Virmani V, Souza CA, Bayanati H, Gupta A, et al. Role of quantitative computed tomography texture analysis in the differentiation of primary lung cancer and granulomatous nodules. *Quant Imaging Med Surg* 2016; 6: 6-15.
28. He L, Huang Y, Ma Z, Liang C, Liang C, Liu Z. Effects of contrast-enhancement, reconstruction slice thickness and convolution kernel on the diagnostic performance of radiomics signature in solitary pulmonary nodule. *Sci Rep* 2016; 6: 34921.
29. Wu W, Ye J, Wang Q, Luo J, Xu S. CT-based radiomics signature for the preoperative discrimination between head and neck squamous cell carcinoma grades. *Front Oncol* 2019; 9: 821.
30. Singh B, Kaur M. An approach for classification of malignant and benign microcalcification clusters. *Sādhanā* 2018; 43: 39.
31. Wang J, Liu X, Dong D, Song J, Xu M, Zang Y, et al. Prediction of malignant and benign of lung tumor using a quantitative radiomic method. Orlando (FL): International Conference of the IEEE Engineering in Medicine and Biology Society; 2016.p. 1272-1275.
32. Tobias Hepp, Matthias Schmid, Olaf Gefeller, Elisabeth Waldmann, Andreas Mayr. Approaches to regularized regression-a comparison between gradient boosting and the lasso. *Methods Inf Med* 2016; 55: 422-430.