

**CT based radiomics signature for phenotyping histopathological subtype in patients with non-small cell lung cancer.**

*Yan Liu, Fadila Zerka, Sylvain Bodard, Mehdi Felfli, Charles Voyton, Alexandre Thinnès, Sebastien Jacques, Antoine Iannessi; MEDIAN Technologies Inc., Valbonne, France; Groupe de Recherche Interdisciplinaire Francophone en Onco-Néphrologie (GRIFON), Paris, France; Median Technologies Inc., Valbonne, France; Median Technologies, Valbonne, France; Centre Antoine Lacasagne, Nice, France*

**Background:** The determination of the histopathology of non-small cell lung cancer (NSCLC) is crucial for guiding the appropriate therapeutic strategy, affecting prognosis and recurrence rates. In addition, targetable oncologic mutations are highly correlated to histological subtypes. Conventional methods such as biopsy or surgical excision are the primary methods for histology determination but are invasive, costly, and have limitations such as sampling error. Computed tomography (CT) scans, widely used for NSCLC diagnosis and follow-up, offer a non-invasive alternative through radiomics-based models that provide comprehensive analysis of the entire tumor and surrounding tissue, improving patient selection and stratification in clinical trials and enhancing the development of molecularly targeted drugs. The goal of this study was to evaluate the ability of a CT-based radiomics model to predict the subtype histopathology of NSCLC patients. **Methods:** A total of 678 patients with advanced-stage NSCLC with at least one measurable lung lesion ( $\geq 10$ mm) were selected from multicenter institutions; 531 were used for training and 147 for independent testing. The test set originates from centers unique from the training cohort. Semi-automatic segmentation of the lung lesions was performed and the segmentations were retrospectively reviewed by an experienced radiologist with  $> 20$  years of experience. A total of 1246 radiomics features were extracted using the Pyradiomics package. Ten Robust features were selected first by removing all features with variances near zero, then by applying Maximum Relevance — Minimum Redundancy to the remaining features. All selected features were standardized before modeling. A support vector machines classifier was trained using a five-fold cross-validation and Random Over-Sampling Examples to classify the histology of NSCLC patients (Squamous vs. Non-Squamous). **Results:** The model trained with the radiomics features achieved a mean area under the curve (AUC) of 0.80 (95% CI,  $\pm 0.05$ ) on the cross-validation and an AUC of 0.77 (95% CI,  $\pm 0.05$ ) on the test set. **Conclusions:** Results showed that the model was able to accurately classify the histology of NSCLC patients as squamous vs. non-squamous in a multicenter setting. It highlights the promise of CT-based radiomics in determining the histopathology or other molecular biomarkers of NSCLC, offering a more efficient, cost-effective, and less invasive alternative to traditional tissue analysis methods. Our findings suggest that this non-invasive method has the potential to improve patient selection and stratification and enhance the development of molecularly targeted drugs. These results underscore the importance of incorporating advanced imaging techniques and machine learning algorithms into oncology practice for better patient outcomes. Research Sponsor: None.