e20515 **Publication Only**

Multicenter evaluation of AI-based CT radiomics for EGFR mutation prediction in NSCLC.

Yan Liu, Sylvain Bodard, Mehdi Felfli, Fadila Zerka, Charles Voyton, Alexandre Thinnes, Sebastien Jacques, Antoine Iannessi; MEDIAN Technologies Inc., Valbonne, France; Hôpital Necker Enfants Malades, Paris, France; Median Technologies Inc., Valbonne, France; Median Technologies, Valbonne, France

Background: Tyrosine kinase inhibitors targeting the epidermal growth factor receptor (EGFR) were the first targeted therapy to enter clinical use for non-small cell lung cancer (NSCLC), thanks to its significant improvement on treatment response and patient survival. Today, tissue sampling is the standard to collect material for testing EGFR mutation status in advanced NSCLC, but has known sampling limitations due to intra-tumoral heterogeneity leading to false negative results. Computed tomography (CT) image based radiomics models offer a non-invasive solution, allowing comprehensive analysis of the entire tumor and surrounding tissue. Radiomics could provide insights during patient longitudinal follow up, or when tissue sample is not accessible is suboptimal. During the past decade, predicting EGFR mutation status using CT radiomics was widely investigated, with promising results. Nevertheless, most studies were single center retrospective studies, without external validation. Therefore, our study aims to investigate the performance of a CT radiomics EGFR mutation prediction model in a multicentral setting with an external cohort. **Methods:** CT images were collected from multicenter institutions and open source databases (https://www.cancerimagingarchive.net/), including a total of 1047 patients with advanced stage NSCLC with at least one measurable lung lesion ($\geq$10mm). Each lung lesions was segmented semi-automatically followed by the validation of an experienced radiologist with $>$ 20 years of experience. The dataset was divided to 2:1 for training and testing, with the test set originating from centers unique from the training cohort. A total of 1246 radiomics features were extracted using the Pyradiomics package. The 10 most robust features were selected first, by removing all features with variances near zero, then by applying Maximum Relevance — Minimum Redundancy (MRMR) to the remaining features. All selected features were standardized before modelling. A support vector machines (SVM) classifier was trained using a five-fold cross validation to classify EGFR mutations status (Wild type vs Mutant). **Results:** The SVM model trained with the radiomics features achieved mean area under curve (AUC) of 0.83 (95% CI, $\pm$0.03) on the cross validation and an AUC of 0.76 (95% CI, $\pm$0.03) on the test set. **Conclusions:** Our multicentric study demonstrated the potential of CT radiomics models in predicting EGFR mutation status in advanced NSCLC, with results similar to those found in current literature. We are optimizing the radiomics model by combining deep learning and clinical features, to improve the accurate, reliable and reproducible performance. AI powered medical image analysis has a bright future as an non-invasive and easy-to-use alternative to identify the expression status of common oncologic genes, serving as predictive biomarker for guiding the targeted therapies of the future. Research Sponsor: Median Technologies.