A Model of Workflow for Radiomics Signature of Stage III Non-Small Cell Lung Cancer (NSCLC) patients (pts)

Alessandra Ferro1, Giulia Pasello1,2, Elena Scaglioni2, Francesca Caimo3, Matilde Costa3, Stefano Frega3, Matteo Sepulcri3, Laura Bononno3, Federico Rea3, Valentina Guarneri1,2, Gissella Gennaro4.

1. Division of Medical Oncology 2, Veneto Institute of Oncology IOV - IRCCS, Padua, Italy; 2. Department of Surgery, Oncology and Gastroenterology, University of Padua, Padua, Italy; 3. Oncologic Radiology Unit, Veneto Institute of Oncology IOV - IRCCS, Padua, Italy; 4. Breast Radiology Unit, Veneto Institute of Oncology - IRCCS, Padua, Italy; 5. Tecnologie Avanzate TA Srl, Udine, Italy; 6. Radiation Therapy Unit, Veneto Institute of Oncology IOV - IRCCS, Padua, Italy; 7. Thoracic Surgery Unit, Department of Cardiac, Thoracic and Vascular Sciences, University of Padua, Padua, Italy.

Abstract

Background

Radiomics is a quantitative approach to medical imaging consisting in the extraction of a large number of features (fts) from diagnostic images that can be associated with tumour pathophysiology and converted into mineable high-dimensional data. We define a model of workflow for extraction and selection of radiomics fts from the baseline computed tomography scan images of stage III NSCLC pts, thus creating a radiomic profile to guide the clinical decision-making process.

Methods

We retrospectively collected data of stage III NSCLC pts referred to Veneto Institute of Oncology from 2012 to 2021. The radiomics pipeline includes (1) the definition of inclusion/exclusion criteria based on image quality and clinical-pathological data, (2) data selection for training and validation cohorts, (3) image segmentation and annotation of target lesions, (4) fts extraction, (5) fts reduction and selection, (6) radiomics signature building. Image segmentation and fts extraction were performed with a commercial software (HealthMyne® Platform), radiomics analysis using the open-source package RadAR (Radiomics Analysis with R) for the statistical programming language R.

Results

On a training cohort of 60 stage III NSCLC pts, we performed the volume segmentation of a maximum of 4 target lesions for each patient. 517 fts were extracted and categorized into morphological, intensity, texture- and filter-based fts. A first stage univariate analysis of the primitive lung target lesion was performed. The feature selection was made by combining different approaches: a principal component analysis to reduce its redundancy, an outlier analysis to exclude its whose distribution showed far outliers, and a further selection on the basis of concordance with clinical data performed with the criterion of minimum redundancy maximum relevance. The correlation of the 10 fts resulting from this pipeline with each relevant clinical datum was then quantitatively evaluated through Spearman’s non-parametric coefficient.

Conclusions

A radiomics workflow for Stage III NSCLC data has been studied. A validation cohort study is warranted in order to combine tumor phenotype characteristics through artificial intelligence applications into predictive and prognostic models.

Contact and Disclosure

E-mail addresses: alessandra.ferro@ioveneto.it; giulia.pasello@ioveneto.it. The authors of this poster have no conflict of interest to declare.

Image Acquisition and Selection

Segmentation

Feature Extraction

Data Integration

Radiomics Signature Building

Radiomics Workflow

Feature selection and dimension reduction methods:

1. an outlier analysis to exclude fts whose distribution showed far outliers;
2. a principal component analysis (PCA) to reduce fts redundancy applied using the R package “RadAR”;
3. on the basis of concordance with clinical data performed with the criterion of minimum redundancy maximum relevance.

Classification of Features

Intensity-based features

N=65

Morphological features

N=54

Feature (matrix-based) features

N=554

Filter-based features

N=54

Kruskal-Wallis test applied to all the radiomic features in order to evaluate which of them can discriminate between clinical dichotomous characteristics.

Correlation map of the 517 extracted features.