Radiogenomics as association between non-invasive imaging features and molecular genomics of lung cancer

Stefania Rizzo¹, Filippo Savoldi², Duccio Rossi², Massimo Bellomi¹,³

¹Division of Radiology, IEO, European Institute of Oncology IRCCS, Milan, IT, Italy; ²Postgraduate School in Radiodiagnosics, ³Department of Oncology and Hemato-Oncology, Università degli Studi di Milano, Milan, IT, Italy

Correspondence to: Stefania Rizzo, MD. Division of Radiology, European Institute of Oncology, IRCCS, Via Ripamonti 435, 20141, Milan, IT, Italy. Email: stefania.rizzo@ieo.it.


Submitted Oct 18, 2018. Accepted for publication Nov 06, 2018. doi: 10.21037/atm.2018.11.17

View this article at: http://dx.doi.org/10.21037/atm.2018.11.17

Cancer is a genetic disease caused by changes to genes controlling the way our cells function, mainly how they grow and divide. Lung cancer is a heterogeneous family of tumors, whose most common type is adenocarcinoma. In 2011, the International Association for the Study of Lung Cancer, the American Thoracic Society and the European Respiratory Society published a multidisciplinary classification of lung adenocarcinoma (1), with further refinements introduced in the World Health Organization (WHO) classification of 2015, integrating genetic and molecular data (2).

At diagnosis, almost all lung cancer patients undergo imaging studies, such as computed tomography (CT) and positron emission tomography (PET), for local staging (3) and to rule out secondary lesions (4-6).

The association between CT descriptors and the pathology of lung cancer has been widely reported in many radiologic-pathologic correlations (7-10). Indeed, the radiological presentation of lung adenocarcinomas includes a broad spectrum of appearances varying from subsolid to solid nodules and masses (1). Along with lesion density, many other descriptors may define the different patterns of lung cancer, such as shape, margins, ground-glass opacity (GGO), cavitation, air bronchogram, and necrosis (11). Although many imaging findings have shown pathologic correlations with adenocarcinoma subtypes and histological patterns (1,12), none have proved strong enough to avoid pathological assessment.

On the one hand, growing evidence supports the concept that a large amount of conventional imaging data not routinely used for reporting can serve to extract information of sufficient depth and complexity to define relationships with underlying tumor genomics (13,14). On the other, recent advances in DNA and RNA sequencing technology have led to an initial understanding of which genomic changes result in the cancer phenotype. These emerging genomic tools, such as analysis of cell-free DNA, RNA and whole exome sequencing, are now available at greater coverage and lower costs, opening further possibilities for patient-tailored lung cancer therapies. This has led to significant changes in the way lung cancer patients are treated in clinical practice. Guided by the presence or absence of specific driver mutations, such as the EGFR mutation or ALK translocation, in advanced stage lung tumors patients may be treated with drugs that specifically target the cells presenting these alterations (15).

Therefore, in the era of precision medicine and targeted therapy, the radiologist must progress from the traditional concept of radiologic-pathologic correlation towards the integration of genomic and phenotypic information provided by new DNA and RNA sequencing technologies and by new ways of analyzing diagnostic imaging modalities.

Radiogenomics is a process designed to extract qualitative and/or quantitative features from volumes of interest, convert them into high-dimensional data, and use them to develop models of diagnosis, prognosis or treatment response (13). When carried out in a robust and structured manner, this process may correlate with large-scale...
molecular information, and there is increasing evidence that genotype-phenotype relationships do scale from genomics to clinical imaging (16). Finding relationships between imaging traits and genomic information is sometimes referred to as creating an association map.

A recent paper by Zhou et al. (17) demonstrated that radiogenic analysis of non-small cell lung cancer (NSCLC) showed multiple associations between semantic image features and metagenes representing canonical molecular pathways, and can result in noninvasive identification of the molecular properties of NSCLC. Specifically, they linked image phenotypes with RNA signatures captured by metagenes, and associated these links with molecular pathways. After evaluating 87 semantic features, they excluded the less frequent ones and then demonstrated the association of the remaining 35 features with the top 10 metagenes. They demonstrated that nodule attenuation and margins were associated with the late cell-cycle genes in their series of 113 NSCLC patients, and a metagene representing the EGF pathway was significantly associated with GGO and irregular nodules or nodules with poorly defined margins. Accordingly, Nair et al. demonstrated that there are several prognostic metagene signatures, the most prognostic one comprising distinct PET-related features, highly correlated with survival also in the external and validation cohorts (18).

Another paper reporting on 212 patients with lung adenocarcinoma surgical stage IA demonstrated a correlation between CT morphology, indicated as pure GGO (39.2%), part-solid nodules (28.8%), or solid nodules (32%), and pathology, indicated as adenocarcinoma in situ (20.8%), minimally invasive adenocarcinoma (29.2%), or invasive adenocarcinoma (50%) with gene mutations (EGFR and KRAS) (19). The authors showed that 36.8% of their cohort harbored an EGF mutation and 8.5% a KRAS mutation, and that a lower GGO component was significantly associated with EGFR and KRAS mutations (19).

A more recent study on 285 NSCLC patients demonstrated radiogenic associations between CT features and the EGF mutation (internal air bronchogram, pleural retraction, small lesion size, and absence of fibrosis), ALK rearrangement (pleural effusion), and the KRAS mutation (round lesion shape and nodules in non-tumor lobes). The authors concluded that the association of these features with significant clinical characteristics, such as female sex and non-smoking for EGFR, young age for ALK, and smoking for KRAS, may suggest which patients are more likely to be mutation carriers (11).

When considering radiogenomics, it is important to choose and incorporate appropriate imaging data. Imaging data may be qualitative (semantic), as in the studies cited above, or quantitative, usually extracted by specific software mainly divided into morphologic and statistical features (20-22).

The creation of an association map may be as simple as representing a relationship between a single image feature and a single molecular or genomic species. Alternatively, it may incorporate complex combinatorial relationships between multiple image features and many molecular or genomic elements, which in combination may define a series of image or molecular phenotypes.

At the molecular pathway-level, gene ontology analysis reveals associations between imaging groups and gene pathways in different types of cancer. For instance, features related to the degree of signal enhancement were associated with the targetable signaling pathways of VEGF and PI3K-Akt and with mTOR signaling, MAPK signaling, focal adhesion and apoptosis. Imaging features indicating necrosis were associated with PI3K-Akt signaling, MAPK signaling, Wnt signaling, and p53 signaling (21).

There is growing evidence that combinations of mutations (rather than a single mutation) are likely responsible for the activation of several pathways/cascades, all leading to different oncogenic endpoints. Therefore, one radiogenomics approach could be to look at gene expression patterns associated with several mutations and imaging features. Another could be to examine broader cancer properties or cancer phenotypes, such as the epithelial-mesenchymal transition. By looking at the specific field of lung cancer radiogenomics, Zhou et al.’s study (17) validated aradiogenic association map linking image phenotypes with RNA signatures captured by metagenes.

Interesting emerging areas of molecular research also focus on novel classes of RNAs, such as microRNAs (miRNAs) and long noncoding RNAs (lncRNAs), which can be evaluated by a number of different transcriptome analyses. Some miRNAs, also defined as small noncoding RNAs (~22 nucleotides), are known to be implicated in lung tumorigenesis with altered expression levels correlating with tumor stage and patient survival. However, no published papers have evaluated the association between imaging features and miRNA expression in lung cancer.

Several lung cancer studies have also shown that long non-coding (lnc)RNAs, also known as noncoding transcripts >200 nucleotides in length, are not translated into proteins, but act as regulatory RNAs, serving as molecular markers for survival, treatment resistance,
and metastases. For example, MALAT-1 may serve as a molecular marker for NSCLC diagnosis, its propensity for metastasis and survival, while CCAT2 may promote invasion and can be considered a biomarker for lymph node metastasis and an independent unfavorable prognostic factor in SCLC patients (12).

Considering the multiple discrete steps needed to extract imaging features (especially quantitative information), each presenting its own challenges (20), and the complexity of the different genomic information that can be used and integrated in radiogenomics studies, it is evident that the process is as composite as it is promising. This accounts for the importance of standardization in radiogenomics studies. Indeed, all radiologists know that it is almost impossible to acquire images according to the same protocol, especially in multicentric studies, because acquisitions are frequently adapted to specific clinical questions (23-25). Nonetheless, there is an increasing need to validate radiomics and radiogenomics studies on independent external cohorts. Therefore, some preliminary image analysis may be required to exclude unreliable and unstable quantitative features. Furthermore, models incorporating multiple levels of validation (e.g., cellular, genetic, protein, clinical, etc.) tend to be more reliable than complex models operating at only one biological level (imaging, or imaging to outcome only) (12).

In conclusion, the era of precision medicine has seen the demise of the concept of radiologic-pathologic correlation, superseded by the rise of radiogenomics. This new direction in cancer research is currently helping scientists understand the multiple-level associations between genomic and phenotypic information encoded in digital clinical images and the underlying clinical and biological correlates, associations, and mechanisms.

Acknowledgements

The English text has been edited by Anne Prudence Collins (Editor and Translator Medical & Scientific Publications).

Footnote

Conflicts of Interest: The authors have no conflicts of interest to declare.

References


Cite this article as: Rizzo S, Savoldi F, Rossi D, Bellomi M. Radiomics as association between non-invasive imaging features and molecular genomics of lung cancer. Ann Transl Med 2018;6(23):447. doi: 10.21037/atm.2018.11.17