Multimodal and integrative analysis of genomics, radiomics and clinical data for the prediction of response to immunotherapy in lung cancer

Nicolas Captier*[1,2,3], Emmanuel Barillot[1], and Irène Buvat[2]

[1]Cancer et génome: Bioinformatique, biostatistiques et épidémiologie dún système complexe – MINES ParisTech - École nationale supérieure des mines de Paris, Université Paris sciences et lettres, Institut Curie [Paris], Institut National de la Santé et de la Recherche Médicale : U900 – France

[2]LITO (U1288) – Université Paris Sud, Université Paris Saclay, Institut National de la Santé et de la Recherche Médicale - INSERM, Institut curie, Orsay – France

[3]Ecole Doctorale Complexité du Vivant – Sorbonne Universite, Université PSL – France

Abstract

Non-small cell lung cancer (NSCLC) is the first cause of cancer-related death in France. NSCLC is diagnosed at a metastatic stage in about 70% of patients. In this situation chemotherapy combined with pembrolizumab is becoming the preferred option even in patients with a tumor with PD-L1 expression $\geq$50%. The survival of patients with metastatic NSCLC has been increasing with such strategies, but only for 45-50% of those patients a response can be demonstrated. Therapeutic decision is thus suboptimal, and there is a critical need for biomarkers for response prediction.

The goal of my PhD is to to develop supervised and unsupervised machine learning methods to identify signatures of the response to immunotherapy in NSCLC, through the integration of genomics, radiomics (extracted from PET and CT images) and clinical data. To that end, I am studying an ongoing prospective cohort of patients with metastatic NSCLC and treated with a combination of chemotherapy and immune checkpoint inhibitors. For each patient, we have access to a wide variety of data (medical images, RNA sequencing of tumor samples, pathological data, clinical data...).

In this talk, I will present different methods we built to analyse the diverse modalities which compose our data and take the best out of it in order to extract relevant information for the prediction of the treatment response. I will focus on the different tools they use to face the wide range of challenges which come with our project : a need for biological interpetation to decipher the mechanisms modulating immune responses, a need for unbiased assesments of the generalizability of our predictive models, or a need for approaches that efficiently combines the advantages and information of each data modality.

*Speaker