

## Détermination du sous-type moléculaire des tumeurs broncho-pulmonaires en imagerie multi-modalité

Nom et coordonnées du responsable du stage : Frédérique Frouin, [frederique.frouin@inserm.fr](mailto:frederique.frouin@inserm.fr)

LITO – Centre de Recherche de l’Institut Curie - Rue de la Chaufferie – Bâtiment 101B - 91400 Orsay

Nature et durée du stage : Stage Recherche de Master 2 ou de fin d’école d’ingénieurs - 6 mois entre février et septembre 2021. Possibilité de télétravail si les conditions de crise sanitaire l’imposent.

Rémunération (estimation mensuelle) : Gratification de stage (environ 550 euros par mois) - Prise en charge de 50% des éventuels frais de transport (Pass Navigo).

### Description :

**Contexte** : La prise en charge thérapeutique des cancers broncho-pulmonaires dépend notamment du sous-type histologique de la tumeur. Dans un certain nombre de cas, les biopsies ne sont pas réalisables ou ne sont pas conclusives, il serait alors important de pouvoir prédire leurs caractéristiques à partir des données existantes, en l’occurrence les données d’imagerie, pour adapter au mieux le traitement. Ce travail utilise les données d’une étude prospective, l’étude IRMomics, conduite par le Pr PY Brillet de l’Hôpital Avicenne.

**Objectifs** : Ce travail vise à identifier le sous-type moléculaire des tumeurs broncho-pulmonaires à partir d’examens IRM ( images pondérées en T1, images pondérées en T2) et d’examens TEP/TDM. Des paramètres cliniques pertinents pour la tâche de discrimination (e.g. le statut tabagique) seront ajoutés aux données d’imagerie. Des facteurs potentiellement confondants tels que le stade TNM de la tumeur seront également pris en compte.

**Méthodes envisagées** : Il s’agira d’améliorer les deux modèles de prédiction de machine learning existants [2,3]. Il s’agit de modèles bi-modalités, l’un utilisant les images IRM pondérées en T1 et celles pondérées en T2 [2], l’autre les images IRM pondérées en T2 et les images TEP [3]. Dans un premier temps, les données issues de patients récemment inclus dans l’étude (environ 50, soit une augmentation de 50% des bases existantes) seront intégrées. Puis, pour pallier le problème des données manquantes (examens non disponibles pour certaines modalités), des approches multi-modèles seront mises en œuvre, chaque modèle prenant en compte les données issues d’une, deux ou trois modalités d’imagerie différentes. Cette approche permet de maximiser les effectifs pour les modèles travaillant sur une seule ou deux modalité(s) et d’accroître ainsi leur robustesse. Pour la validation, une vingtaine de nouveaux patients seront disponibles. Chacun sera testé en lui appliquant l’ensemble des modèles de prédiction pour lesquels les données d’imagerie sont disponibles. Différents scores (pondérés éventuellement par le degré de confiance attribué à chaque modèle) combinant les résultats seront finalement proposés pour évaluer la pertinence de l’approche multi-modèle. L’étude sera également faite en intégrant les données cliniques potentiellement intéressantes aux données d’imagerie.

### **Références bibliographiques :**

[1] Lacroix et al. Correction for Magnetic Field Inhomogeneities and Normalization of Voxel Values Are Needed to Better Reveal the Potential of MR Radiomic Features in Lung Cancer. *Front Oncol.* 2020;10:43.

[2] Lacroix et al. Complementarity of T2-weighted and contrast-enhanced T1-weighted MR radiomic features for characterization of lung cancers. *MAGMA.* 2020;33(Suppl1) :S109-S110.

[3] Orhac et al. Complementarity of PET and MR radiomic features for the characterization of lung cancers. *EJNMMI* 2020;47(Suppl1):S332.

Compétences requises : Expérience de programmation en langage Python et en langage R. Expérience en analyse d’images. Bases solides en statistique et en machine learning.