

Application de machine learning en imagerie médicale : validation de l'approche Radiomics sur une population de patients traités par radio(chimio)thérapie aux Cliniques universitaires Saint-Luc pour un cancer du poumon non à petites cellules

Madeleine Scrivener

Promoteur : Pr. Xavier Geets



INTRODUCTION

De plus en plus de progrès sont réalisés dans le domaine de l'imagerie médicale et plus particulièrement dans la conversion d'images médicales en données utilisables. Radiomics est une méthode d'extraction à haut débit de caractéristiques quantitatives d'images médicales (CQI) (regardez <https://youtu.be/Tq980GEVP0Y> et consultez le site www.radiomics.world). Ces « CQI » peuvent être divisées en quatre groupes en fonction de la caractéristique de la tumeur qu'elles décrivent : son intensité, sa forme, sa texture ou ses caractéristiques à travers des filtres (dit « wavelets »). Ici nous présentons une analyse radiomique des 18 026 « CQI » extraites d'images 4D CT pré-traitement d'une cohorte de 44 patients traités par radio-chimiothérapie pour un cancer du poumon non à petites cellules aux Cliniques universitaires Saint-Luc. Une signature radiomique est créée en utilisant des algorithmes d'apprentissage automatique qui présélectionnent un petit nombre de CQI sur base de leur corrélation avec le résultat étudié (survie, types histologique...). Cette signature est construite en analysant les CQI d'une première cohorte de patients et doit ensuite être testée et validée sur une cohorte externe afin de voir si la signature Radiomique est robuste.

HYPOTHÈSES

Premièrement, nous avons fait l'hypothèse que les CQI extraites à partir de CT thoraciques pré-traitement contiennent des informations sur le type histologique de la tumeur, nous permettant de créer une nouvelle signature différenciant les adénocarcinomes des carcinomes épithéliaux. Ensuite nous avons fait l'hypothèse que nous pouvions valider deux signatures radiomique à visée pronostique précédemment publiées (*Aerts et al* et *Tunali et al*) sur la cohorte de St-Luc.

MATÉRIEL ET MÉTHODE

Notre étude a été divisée en quatre parties. **Premièrement**, nous avons créé une nouvelle signature Radiomique qui différencie les

patients avec un adénocarcinome histologiquement confirmé de ceux qui ont une tumeur d'un autre type histologique sur base de leurs CQI. Cette signature a été créée à partir d'une cohorte open-source de 422 patients traités à Maastricht (MAASTRO Lung). Nous avons ensuite validé cette signature sur la cohorte de 44 patients de St-Luc et sur une cohorte de 99 patients traités dans le service de radiothérapie à l'Université de Californie à San Francisco (UCSF). Pour la **deuxième partie**, nous avons validé une signature radiomique à visée pronostique publiée par *Aerts et al.* [3] en 2013 dans *Nature Communications* sur la cohorte traitée à St-Luc. Cette signature utilise quatre CQI pour diviser la cohorte de patients en deux groupes, un de meilleur pronostique et l'autre avec un pronostique plus réservé. Et **troisièmement**, nous avons validé une autre signature pronostique développée au Moffitt cancer center en Floride par *Tunali et al.* [61] et publiée dans *Oncotarget* en 2017. Cette signature utilise deux CQI (le « degré radial (RG) » et la « déviation radiale (RD) ») extraites des images de patients pour les répartir les patients en un groupe de meilleur pronostique avec une tumeur plus indolente et un groupe qui présentent des tumeurs plus agressives et qui ont donc un pronostic plus réservé. **Dernièrement**, nous avons évalué la qualité de la méthodologie de cette étude à l'aide de deux scores internationaux : le Radiomics Quality Score (RQS) et les recommandations TRIPOD.

RÉSULTATS

La première partie de l'étude a mis en évidence des résultats significatifs dans la différenciation de patients sur base du type histologique de leur tumeur, création de la signature (AUC : 0.93) et validation (AUC : 0.82). Les résultats de la validation des deux signatures à visée pronostique étaient statistiquement non significatifs mais prometteurs et compatibles avec nos hypothèses car il y a une distincte séparation des deux groupes sur les courbes de Kaplan-Meier. Ceci souligne le potentiel des CQI, qui fournissent des informations supplémentaires de façon non-invasive sur la prolifération tumorale, le pronostic du patient et le type histologique sur base de leurs CT pré-traitement. Cette approche pourrait avoir un réel impact clinique mais nécessite pour cela des études prospectives supplémentaires qui sont déjà en cours.

MOTS-CLÉS

Imagerie médicale, machine learning, Radiomics, modèles prédictifs, caractéristiques quantitatives d'images, cancer du poumon non à petites cellules, radio-chimiothérapie, histologie

Machine learning applied to imaging: validation of the radiomics approach in a population of non-small-cell lung cancer patients treated by (chemo-)radiotherapy at the Cliniques universitaires Saint-Luc

INTRODUCTION

An increasing number of advances have been achieved in the field of medical imaging, namely the conversion of standard-of-care images into mineable data. Radiomics refers to the high-throughput extraction of quantitative image features from medical images (watch the video on <https://youtu.be/Tq980GEVPOY> and visit www.radiomics.world). These image features can be divided into four groups depending on the tumor characteristic they describe: tumor intensity, tumor shape, tumor texture, or wavelets. This study consisted in a radiomics analysis of 18'026 features extracted from standard-of-care, pretreatment, 4D computed tomography (CT) images from a cohort of 44 non-small-cell lung cancer (NSCLC) patients treated with chemoradiotherapy at the Cliniques Universitaires St-Luc. A radiomics signature was created using machine learning algorithms in order to preselect a small group of radiomics features based on their correlation with the studied endpoint (survival, histological types, etc.). The signature was created by analyzing the features of a patient cohort, and its robustness was then tested and validated on an external independent dataset.

AIMS AND OBJECTIVES

First, we hypothesized that adenocarcinomas could be differentiated from squamous cell carcinomas using a new radiomics signature. This would show that the radiomics features extracted from pretreatment CTs contain information about the tumor's histological type. We further hypothesized that we could validate two previously published prognostic radiomics signatures (Aerts *et al.* and Tunali *et al.*) on the St-Luc cohort.

MATERIAL AND METHOD

Our study was divided into four parts. In the **first part**, we created a radiomics signature in order to differentiate histologically

confirmed adenocarcinoma from other histological types in an open-source cohort of 422 NSCLC patients from the Maastricht University. We then validated the signature in the cohort of 44 patients treated at St-Luc and in another cohort of 99 NSCLC patients treated in the Department of Radiation Oncology at the University of California, San Francisco (UCSF). In the **second part**, we validated the Aerts *et al.* prognostic signature published in 2013 in *Nature Communications* in the St-Luc cohort. This signature uses four radiomics features describing the tumor so as to divide the patients into two different prognostic groups. **Thirdly**, we validated the published signature of Tunali *et al.* in the St-Luc cohort. This signature is also a prognostic signature that uses only two radiomics features, namely radial gradient (RG) and radial deviation (RD), to divide the patients into a group with better prognosis and more indolent tumor and another group with more aggressive disease and poorer prognosis. **Lastly**, we used two different international scoring systems to evaluate the methodology of our approach: the Radiomics Quality Score (RQS) and the TRIPOD recommendations.

RESULTS

The first part of the study showed significant results for training (AUC: 0.93) and validation (AUC: 0.82) of the histological radiomics signature. The two prognostic radiomics signatures showed promising, though not statistically significant, results, thereby highlighting the potential of radiomic features and machine learning for non-invasively providing additional information on the tumor's proliferation and histological type. This could eventually be used in routine practice, thus improving therapeutic decision-making and reducing complications associated with invasive diagnostic procedures. This study will be continued on a larger cohort of St-Luc patients.

KEY WORDS

Medical imaging, machine learning, radiomics, predictive models, image quantitative characteristics, non-small-cell lung cancer, chemo-radiotherapy, histology

AFFILIATIONS

Cliniques universitaires Saint-Luc, Service de Radiothérapie, avec l'aide du GROW-School for Oncology and Developmental Biology, Maastricht University Medical Centre, Maastricht, The Netherlands