Contributions to radiomics development for precision medicine in neuro-oncology

Advisors Frédérique Frouin, CR Inserm – Irène Buvat, DR CNRS
Laboratoire Imagerie Moléculaire In Vivo (IMIV), UMR 1023 Inserm/CEA/Université Paris Sud - ERL 9218 CNRS. CEA/I2BM/SHFJ 4 place du Général Leclerc 91400 Orsay
Contact: frederique.frouin@inserm.fr

Abstract

Radiomics is an emerging field that aims at building a relevant statistical model from a large number of features extracted from medical imaging data (possibly combined with clinical or genomic data) to assist diagnosis, prognosis and therapy monitoring.
This field integrates various methodological skills including medical physics, medical image processing, data analysis, machine learning and biostatistics.
Radiomics is based on a robust extraction of features from medical images. It requires a succession of image processing steps: artifact correction, segmentation of tumor and reference regions, standardization of image signal intensity, tumor feature extraction. Using these features, a statistical model is designed based on machine learning algorithms, which have to be tuned according to the clinical or biological question and to the a priori knowledge that is available. Once developed, the model has to be tested on independent data, in order to assess its contribution to personalized patient treatment of. All these steps raise specific issues that have not been solved yet.
The objective of the current project is to investigate the impact of the choices involved in the different image processing steps and in the machine learning algorithms on the quality of the predictive models, to try to optimize the radiomics methodology. Multi-parametric magnetic resonance imaging data acquired as part of a pediatric neuro-oncological research protocol currently performed in Gustave-Roussy and Necker hospital will be used. The clinical objective is to identify some imaging features that could predict response to radiotherapy and other features that could predict survival rate when embedded in a radiomics approach.
Expected results include significant contributions to the development of radiomics as well as some demonstration of its added value for precision medicine. Specific care will be devoted to the developments of tools and methods applicable in clinical research protocols.

Detailed subject

Context

The recent development of radiomics (Lambin et al. 2012, Kumar et al. 2012, Aerts et al. 2014) generates great expectation for precision medicine in oncology. Indeed, the high number of data acquired cannot be fully exploited in clinical routine, due to a lack of appropriate tools. To overcome this limitation, radiomics assumes that a comprehensive analysis of a large number of quantitative features extracted from multimodal medical images could improve prognosis or patient management with respect to conventional markers, such as tumor dimensions as measured from CT or MRI or tumor metabolic activity assessed by PET. Before being applied to prognostic studies, an intermediate objective consists in establishing a link between image features and molecular markers related to the biology of tumors. For instance, a systematic study of several dozens of imaging features has established a connection between specific molecular processes and a limited number of quantitative imaging features extracted from a MRI genetics study in Glioblastoma multiforme (Gevaert et al. 2014). Some recent works have shown the potential of characterizing tumor heterogeneity using image texture analysis (Buvat et al. 2015). However, as shown by Orlhac et al. (2014), a relevant texture analysis in PET requires standardization of the acquisition protocol, of the reconstruction and of the index calculation. In case of multiple time point imaging, which is commonly used for therapy monitoring, specific care must be paid when comparing tumor data of the same subject at the voxel level, including an appropriate realignment of the different scans (Tacchella et al. 2014).
For this project, multi-parametric MRI data obtained from a pediatric protocol in neuro-oncology involving Gustave-Roussy and Necker hospital will be analyzed. The clinical model under study is the diffuse intrinsic pontine glioma, one of the most aggressive pediatric brain cancer for which radiotherapy is currently the only treatment available to patients. MRI data include T1 weighted and FLAIR images, perfusion images, diffusion weighted images and spectroscopic data. A retrospective database of 30 patients acquired on a 1.5 T scanner will be available at the beginning of the project. A prospective database including 100 patients will be accumulated on the newly acquired 3T MR scanner at Necker Hospital (BIOMEDE protocol, PHRC).

Thesis objectives

From the methodological point of view, this work will address a robust feature extraction from multi-parametric MRI and will design well-suited machine learning methods. First, this study will address the impact of image corrections, tumor segmentation and MRI standardization on the quality of feature extraction (Gevaert et al. 2014, Vallières et al. 2015). Regarding the subsequent statistical analyses, due to similarities between image and genetics data, some algorithms already used for genetic data will be tested and adapted. Two clinical objectives will be pursued: 1) defining which combination of MRI parameters could be associated with the molecular variants of patients (histone H3 variant genes); 2) defining which combination of MRI parameters could predict radiotherapy outcome, early after the treatment.

Scientific program

First, MRI data acquired just before biopsy will be analyzed to extract predictive features of the different tumor subtypes (H3K27M mutation). The region of biopsy will be defined using a posteriori MR data. To achieve this goal, the different scans associated with each patient will be co-registered, accounting for magnetic field heterogeneities (this effect is expected to be increased at 3T). The whole tumor and the biopsy region will be segmented semi-automatically on T2 weighted images, then tumor edema and necrotic sub-regions will be identified from the multiple imaging data obtained for each patient. This will yield an intermediate characterization of the tumor heterogeneity between the global (tumor region) and the voxel-based analysis.

All these ROIs will be used for subsequent analysis: conventional shape descriptors (size, eccentricity, compactness, etc.), relative positioning of these different sub-regions will be measured as well as signal intensity values, histograms and texture parameters in various MR images (T1-, T2-, FLAIR weighted images, parametric images derived from Gadolinium and Arterial Spin Labeling based perfusion studies and multi-b, ADC derived diffusion images). Reference values will also be calculated in distant healthy regions (white matter estimated in the brain regions less affected by radiotherapy). Univariate and correlation analyses will be conducted to select the most relevant features.

The feature robustness will be assessed with respect to various parameters, including segmentation, use of absolute or normalized values, and discretization used for texture analysis. Differences in feature values between 1.5 T and 3T scanners will be also studied, opening the way towards the integration of multi-centric studies.

Multivariate analysis will be then developed, testing different machine learning algorithms applied on a selection of features. Another option that will be investigated is the integration of all features in the analysis associated with a parsimony constraint in the underlying model.

To predict the radiotherapy efficacy, several time points might be used. Thus the specific handling of longitudinal features in radiomics will be tackled. Due to preliminary results obtained on a small group of patients, the predictive value of changes in perfusion and diffusion parameters over time will be first investigated.

Expected outcomes

Expected results concern the development of robust methods for radiomics and the proof of its added value compared to current diagnostic and prognostic methods. The methods will be designed with a view to enable practical application of the radiomic approach to research protocols. Indeed, we aim at providing a first practical experiment to joint analysis of MR multi-parametric images that could ultimately be extended to the analysis of multi-parametric PET/MR data acquired on our new PET/MR SIGNA GE PET scanner.

References

(Aerts et al. 2014) Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach, Nature Communications 5:4006:1-8.

(Buvat et al. 2015) Tumor Texture Analysis in PET: Where Do We Stand? J Nucl Med 56:1642-4.

(Gevaert et al. 2014) Glioblastoma multiforme: exploratory radiogenomic analysis by using quantitative image features. Radiology 273:168-74.

(Kumar et al. 2012) Radiomics: the process and the challenges. Magn Reson Imaging 30:1234-48.

(Lambin et al. 2012) Radiomics: extracting more information from medical images using advanced feature analysis Eur J Cancer 48:441-6.

(Orlhac et al. 2014) Tumor texture analysis in 18F-FDG PET: relationships between texture parameters, histogram indices, standardized uptake values, metabolic volumes, and total lesion glycolysis. J Nucl Med 55:414-22.

(Tacchella et al. 2014) An efficient strategy based on an individualized selection of registration methods. Application to the coregistration of MR and SPECT images in neuro-oncology. Phys Med Biol 59:6997-7011.

(Vallières 2015) A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities. Phys Med Biol 60:5471-96.