International Doctoral School
Medical University of Lodz

BTM

Diagnostic, prognostic and predictive capabilities of radiomic analysis in oncology

Szymon Grabia M.Sc. Eng.,
Supervisors: Prof. Wojciech Fendler M.D., Ph.D.;
Urszula Smyczyńska M.Sc., Ph.D.
Department of Biostatistics and Translational Medicine

Introduction

Radiotherapy (RT) is known to lead to both immune stimulatory and suppressive effects, the latter of which can cause such side effects as radiation-induced lymphopenia (RIL). RIL can lead to an increased risk of infections, a compromised immune system, and a less effective response to cancer treatment. It was reported as a negative prognostic factor in patients treated with RT for variety of tumors including lung, brain, pancreatic, and cervical cancers. Thus, early identification of high risk of RIL would help in replanning RT in such a way that RIL could be avoided. The approach for identifying biomarkers of radiotherapy complications that I am investigating utilizes the high-throughput capabilities of radiomics. It allows for the extraction of features, especially those invisible to the human eye, from medical images in a fast, unbiased manner and analyzing them using data mining methods. Since Computed Tomography (CT) reflects only patients' anatomy, but not organs' exposure to radiation, my idea was to add to the analysis radiomic features extracted from a 3D dose distribution (RTDOSE), generated from treatment planning, done before RT. As seen in Fig. 1., even though applying high doses of radiation to a heart should usually be avoided, it is not always possible and study of the patterns of radiation in such regions of interest could find biomarkers of RIL.

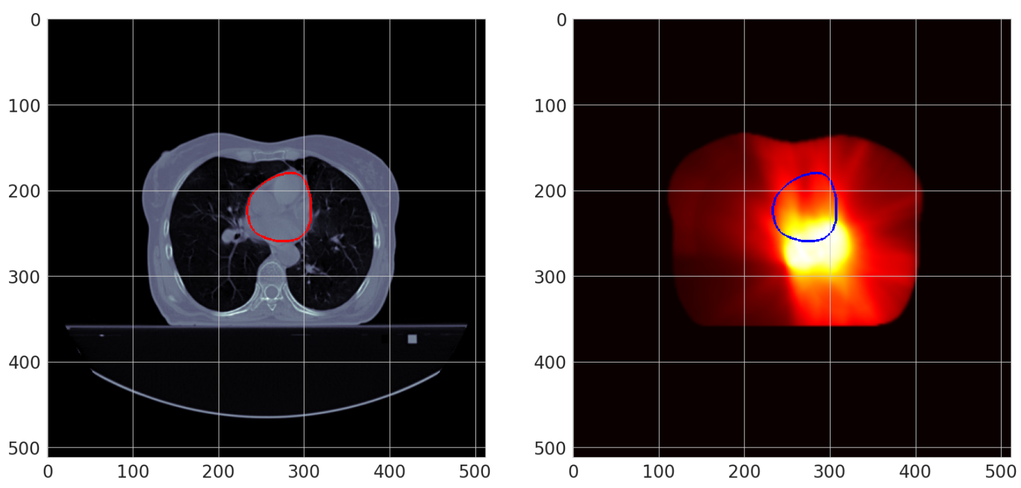

Figure 1. Sample CT and RTDOSE slices of thorax region with heart outlined

Methods

The population for my study consisted of patients (Table 1), who have undergone radical radiotherapy for non-small cell lung cancer and whose planning process was performed manually by doctors based on planning CT scans. The occurrence of RIL was assessed based on the lymphocyte count measured during radiotherapy. As a first step, I constructed an automated pipeline for transforming the raw medical images into corresponding sets of radiomic features, including extraction of DICOM metadata, creation of organ masks based on RTSTRUCT (outlines of organs created by radiotherapists), scaling of RTDOSE to match those mask volumes and extraction of radiomic features from both CT and RTDOSE modalities using PyRadiomics Python package. Then, after appropriate processing of feature values, I developed several machine learning models to select the best classifier for RIL prediction. The organ chosen for the investigation was a heart, because it is always delineated manually for lung cancer RT as it is considered as an organ at risk that should receive the lowest possible radiation dose.

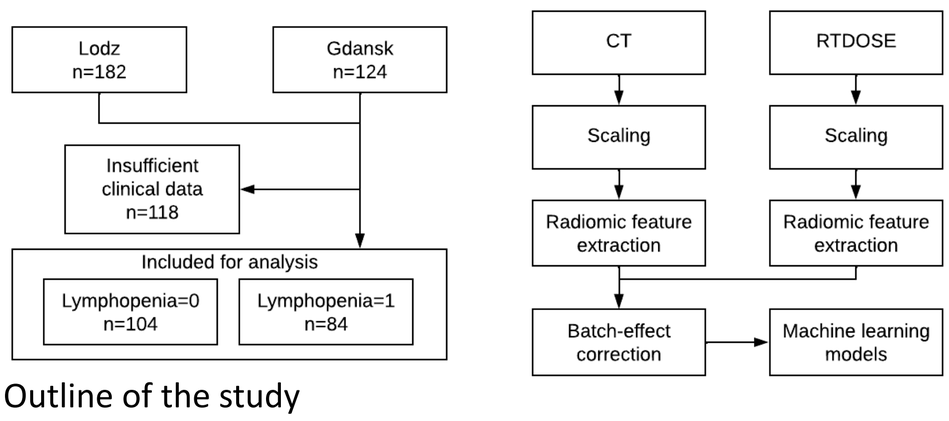

Figure 2. Outline of the study

Table 1. Patient characteristics

		Lymphopenia		
		NO	YES	pvalue
Age	(mean)	65.6256	65.636	0.9920
Sex	F	32	26	0.8752
	M	62	53	
Stage	1	2	1	0.7607
	2	6	7	
	3	86	71	

		Lymphopenia		
		NO	YES	pvalue
TNM - T	1	11	9	0.7910
	2	32	20	
	3	30	29	
	4	19	19	
TNM - N	0	8	7	0.4651
	1	14	18	
	2	66	47	
	3	6	7	
TNM - M	0	94	79	NA

Results

In addition to the data presented last year, I managed to collect further images from a center in Lodz. Thanks to the cooperation with the Medical University of Gdansk, I was also able to acquire additional medical scans for patients fulfilling my recruitment criteria from this center. I managed to collect DICOM sets for 182 patients from Lodz and 124 from Gdansk for a total of 306 sets. In each of these sets, at least 1 series for each patient was available, consisting of CT, RTDOSE, RTPLAN and RTSTRUCT modalities. After removing records with insufficient clinical data, 188 patients were included, 84 with RIL, 104 without. The preprocessing and feature extraction steps yielded 1559 radiomic features from CT and 1567 from RTDOSE for each patient, which were viable for further analysis. Since the range of pixel values for each of those modalities was notably different, some features also consequently had inconsistent ranges of values. Thus, radiomic features had to be first scaled within their modalities, then merged. When performing PCA, the main differentiating factor was the patient's assignment to a center, so the use of a batch correction through ComBat was required. Finally, a logistic regression model was chosen, as it performed the best, with area under the ROC curve (AUC) equal to 0.7743 (Fig. 2).

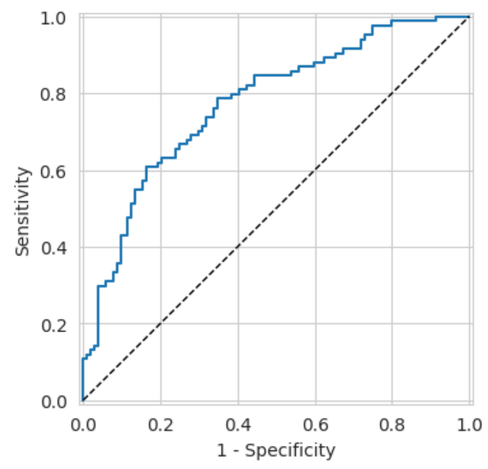

Figure 3. ROC curve for selected logistic regression model

Conclusions

Successful creation of automated pipelines for processing high volumes of DICOM images allowed me to reliably handle all the supplied data from two medical centers and calculate corresponding radiomic features. My previous investigations revealed that radiomics features obtained from CT scans can be used as an acceptable predictor of complications after radiotherapy, however they did not significantly outperform commonly used state-of-art normal tissue complication probability (NTCP) models. The addition of radiomic features extracted from RTDOSE distributions seemed to be a reasonable extension, as RTDOSE represents the actual projection of a radiotherapy plan. The performance of a model utilizing combined CT and RTDOSE radiomic features was adequate. Machine learning models based on radiomic features are still prospectively useful, since they provide a comprehensive and noninvasive outlook on patients' condition, simultaneously being a mostly undiscovered field. Lastly, analogous analysis will be performed for other organs commonly associated with developing RIL.

References

1. Van Griethuysen, J. J., Fedorov, A., Parmar, C., Hosny, A., Aucoin, N., Narayan, V., Beets-Tan, R. G. H., Fillion-Robin, J., Pieper, S., & Aerts, H. J. (2017). Computational Radiomics system to decode the radiographic phenotype. Cancer Research, 77(21), e104–e107. https://doi.org/10.1158/0008-5472.can-17-0339