PhD Thesis Title: Radiation therapy outcome prediction using statistical correlations & deep learning

Author: André Diamant
Email: andre.diamantboustead@mail.mcgill.ca
Institution: McGill University
Supervisor: Dr. Jan Seuntjens & Dr. Issam El Naqa
Graduation Date: Spring 2020
Available Online: https://escholarship.mcgill.ca/concern/theses/2n49t648k?locale=en

ABSTRACT:

Prognosis after cancer treatment is a constant concern for physicians, patients and their surrounding friends and family. This is one of the reasons that treatment outcomes prediction is such a critical field of research. The sheer magnitude of data generated within a typical radiation oncology clinic each year facilitates the development and eventual validation of predictive and prognostic models. Furthermore, the technological advances driven by data science have enabled the usage of advanced machine learning techniques which can far exceed the performance of previously used conventional techniques.

Most cancer patients follow a standard radiation oncology workflow, which among other things includes medical imaging and the creation of a radiation therapy treatment plan. As these sorts of data are (in theory) present for every patient, they are ideal variables to input into a predictive model. The goal of this thesis was to investigate these two types of pre-treatment input data (diagnostic imaging and dosimetric data) along with patient characteristics to identify associations and create models capable of predicting a cancer patient's treatment response following radiation therapy.

The first objective was to investigate dose-volume metrics as predictors of clinical outcomes in a cohort of 422 non-small cell lung cancer (NSCLC) patients who received stereotactic body radiation therapy (SBRT). A correlation between the dose delivered to the region outside the tumor and the occurrence of distant metastasis was revealed. In particular, patients who received above a certain threshold dose were shown to have significantly reduced distant metastasis recurrence rates compared to the rest of the population. This was first shown on 217 patients all of whom were treated with conventional SBRT treatment modalities. Next, a similar analysis was done on 205 patients who were treated with a robotic arm linear accelerator (CyberKnife). It was found that the CyberKnife cohort had both superior distant control and local control, suggesting that under current prescription practices, CyberKnife, as a delivery device, could be superior for treating NSCLC patients with SBRT.

The second objective of this thesis was to investigate the usage of a deep learning framework applied to raw medical imaging data in order to predict the overall prognosis of head & neck cancer patients post-radiation therapy. A de novo architecture was built incorporating CT images, resulting in comparable performance to a state-of-the-art study. Furthermore, our model was shown to recognize imaging features (`radiomics') previously shown to be predictive without being explicitly presented with their definition. The final portion of this work was the development of a multi-modal deep learning framework which incorporated CT & PET images along with clinical information. This was compared to the previous architecture built, showing substantial increase in prediction performance for both overall survival and local recurrence. It was also shown to function in the presence of missing data, a common occurrence within the medical landscape.

This work demonstrates that pre-treatment prediction of a cancer patient's post-radiation therapy outcomes is possible by learning correlations and building models from readily available data. Future efforts should be put towards data sharing & data curation to enable the creation and validation of models that eventually can be used in the clinic. Ultimately, predictive models should evolve into generative models whereupon one's treatment could be automatically created with the explicit intention of statistically optimizing that patient's outcomes.

References to author publications that relate specifically to the dissertation:

1. **Diamant, A**., Chatterjee, A., Faria, S., Naqa, I. E., Bahig, H., Filion, E., Robinson, C., Al-Halabi, H., & Seuntjens, J. (2018). "Can dose outside the PTV influence the risk of distant metastases in stage I lung cancer patients treated with stereotactic body radiotherapy (SBRT)?" *Radiotherapy and Oncology*. Volume 128(3), P513-519. https://doi.org/10.1016/j.radonc.2018.05.012
2. **Diamant, A.,** Chatterjee, A., Vallières, M., Shenouda, G., & Seuntjens, J. (2019). "Deep learning in head & neck cancer outcome prediction." *Scientific Reports 9*, 2764. https://doi.org/10.1038/s41598-019-39206-1
3. **Diamant, A.,** Heng, V. J., Chatterjee, A., Faria, S., Bahig, H., Filion, E., Doucet, R., Khosrow-Khavar, F., Naqa, I. El, & Seuntjens, J. (2020). "Comparing local control and distant metastasis in NSCLC patients between CyberKnife and conventional SBRT." *Radiotherapy and Oncology*, Volume 144, P201–208. https://doi.org/10.1016/j.radonc.2020.01.017