Defining the biological and clinical basis of radiomics

Summary

SUMMARY

Radiomics is a promising field that aims at leveraging existing medical imaging data with advanced computational methods to provide novel quantitative and actionable data for clinical decision support. Notably, radiomic approaches enable integration of these previously untapped sources of information in a noninvasive, objective, reproducible, and cost-effective way. Numerous studies have suggested prognostic and predictive associtions of radiomics and clinical endpoints, including overall survival. However, the molecular biology underlying these associations were largely unknown. To this end, this dissertation contributed in unraveling the connections between radiomic, molecular, and clinical data. Furthermore, these results were put in context to radiomic predictions for targeted therapies, as well as in context of optimizing these predictions with advanced machine learning methods.

This thesis was organized in five parts, where Part 1 introduced and outlined the objectives, Parts 2 to 4 contained the main research results, and Part 5 discussed these results. We briefly summarize the content of the thesis below.

Part 1: Introduction

In **Chapter 1** we introduced the concept of radiomics and explained the implications of radiomics for precision medicine in cancer. In this chapter, we outlined the objectives of this thesis to uncover the underlying biology of radiomics and therefore reason how radiomics can be used as imaging biomarkers for targeted therapies. We motivated these objectives by the fact that radiomics has been suggested to contain prognostic and predictive value, but that biological reasoning of this has not been achieved yet.

Part 2: Radiomics and its Underlying Biology

Chapter 2 presented a broad radiogenomic study that revealed mechanistic connections of a large set of radiomic CT features, molecular pathways, and clinical factors in lung cancer. Importantly, all associations were validated in independent data. Additionally, key associations were also validated biologically. Furthermore, we for the first time demonstrated that radiomics provides complementary prognostic value to traditional genetic and clinical predictors.

In **Chapter 3** and **Chapter 4** we extended the results from **Chapter 2** to associations between volumetric phenotype features, molecular pathways, and somatic gene mutations in brain cancer. In this way, we complement **Chapter 2** by sheding light on how imaging-genomic connections behave in another aggressive cancer type assessed by an alternative imaging modality, namely glioblastoma and MRI, respectively.

Part 3: Radiomics for Targeted Therapies

Understanding genotype-phenotype interactions in tumors as aimed for in Chapters 2-4 allows reasoning about radiomic phenotype predictors for targeted therapies that exploit genetic properties of tumors. Therefore, in **Chapters 5 and 6** we aimed at developing novel imaging biomarkers to identify patients who would respond best to bevacizumab and gefitinib in brain and lung cancer, respectively. In both studies, we leveraged prospectively acquired data of previously published clinical phase II trials.

Particularly, in **Chapter 5** we achieved to develop (and validate) radiomic biomarkers for patients with recurrent glioblastoma that stratified overall survival, as well as progression-free survival, at multiple timepoints including pre-treatment and follow-up at six and twelve weeks post-treatment initiation of bevacizumab. This study strongly suggests that radiomics can be used to develop predictive biomarkers for stratifying this patient population prior to treatment decision, which could be crucial as recent phase III clinical studies failed to suggest improvement of overall survival due to absence of effective biomarkers.

In **Chapter 6** we explored radiomic predictors for gefitinib, an EGFR inhibitor, by assessing predictive value of radiomics for EGFR mutations. Similar to Chapter 5, we investigated imaging data pre- and post-treatment, as well as the delta change of features values between those timepoints. While we observed predictive value of radiomic features at baseline (i.e., before treatment), traditional volumetric features such as tumor volume and diameter showed no predictive value at baseline.

Part 4: Prognostic Value of Radiomic Machine Learning

As we have observed from the previous chapters, radiomics is moving towards a data science, critically dependent on efficient machine learning methods. Therefore, in our final research **Chapters 7-9** we investigated applicability of a wide range of machine learning algorithms to optimize radiomic predictions.

In **Chapter 7** we present a study that evaluated a host of potential machine learning algorithms for radiomic prognostication. Specificially, we investigated 14 feature selection and 12 classification methods, and compared their performance in predicting overall survival at two years. With this study, we aimed at providing recommendations as to which popular machine learning methods are the most promising ones for the development of radiomic prognostic biomarkers.

Finally, wih **Chapter 8-9** we used unsupervised feature selection to identify a canonical set of radiomic features that could potentially predict pathological response in lung cancer following neoadjuvant chemoradiation. In **Chapter 8** we used a similar methodology to

derive features that predicted the development of distant metastasis, while optimizing the predictive model with supervised feature selection.

Part 5: Discussion and Future Perspectives

This dissertation concluded with a discussion and outlook in **Chapter 10**. Our research uncovered a large proportion of the molecular biology underlying radiomics, results that are necessary to justify translation of radiomics into clinical applications. We now gained better understanding about the high prognostic potential of combining radiomics with genomics and clinical records, as well as the implications for targeted therapies. Furthermore, we propose advance machine learning algorithms to increase prognostic and predictive performance of radiomic models. In conclusion, our work demonstrates the significance of systematically analyzing routine medical images with radiomics, as well as its integration with molecular and clinical data, to develop predictive tools for cancer diagnostics.