American Society of Clinical Oncology – 2022 Abstract

Incorporation of intrapatient response heterogeneity using ^{18}F-NaF PET/CT imaging improves outcome prediction models for metastatic prostate cancer patients

Timothy G. Perk, Ph.D.[1]*, Stephen S F Yip, Ph.D.[1], Amy J. Weisman, Ph.D.[1], Sean S. Houshmandi, Ph.D.[1], Elisabeth I. Heath M.D.[2], Michael J. Morris M.D.[3], Ravi Madan, M.D.[4], Douglas McNeel M.D., Ph.D.[5], Andrea B. Apolo, M.D.[4], Christos E. Kyriakopoulos, M.D.[5], Glenn Liu, M.D.[1, 5], Robert Jeraj, Ph.D.[1,5]
[1]AIQ Solutions – Madison, WI ,[2]Karmanos Cancer Institute, Wayne State University School of Medicine, Detroit, MI,[3]Memorial Sloan Kettering Cancer Center, New York, NY,[4]National Cancer Institute, Bethesda, MD, [5]University of Wisconsin – Madison, WI

Background: Quantitative ^{18}F-NaF PET/CT imaging metrics have been shown to be prognostic in metastatic prostate cancer (mPC) patients. However, previous studies have shown conflicting results in which metrics could be prognostic. This study investigates if current methods from literature generalize to external datasets and explores which combination of features are necessary to for survival models to generalize across datasets.

Methods: Imaging and progression-free survival (PFS) data from 118 patients with mPC from four separate prospective clinical trials were gathered retrospectively. Patients received ^{18}F-NaF PET/CT imaging at baseline and at follow-up, between eight and thirteen weeks. TRAQinform IQ technology (AIQ Solutions) was used to identify, segment, and track individual lesions from baseline to follow-up. Eighty-four imaging features were extracted from each patient and sorted into baseline, follow-up, response, patient-level (no inter-lesion comparison), and intrapatient heterogeneity (comparisons between lesions). The data was split into two training and testing sets, 44 patients from one study and 73 patients from the remaining 3 studies. As they can utilize large number of inputs without overfitting, random survival forest (RSF) models were chosen to evaluate performance of feature sets in predicting PFS. Different combinations of features were used as inputs to RSF models to compare single timepoint features with response features and patient-level features with intrapatient heterogeneity features. The performance of the RSF models, together with other methods identified in literature, were evaluated in each dataset using Kaplan-Meier analysis for categorical variables and the c-index for continuous variables.

Results: No patient-level imaging features highlighted by literature displayed significant association to PFS across all four clinical trials (c-index < 0.62 in at least one dataset). Other criteria from literature did not generalize across all datasets ($P>0.05$). The RSF model trained with all features had high c-indices in all four datasets (range: 0.66-0.80). RSF models built with response features (min: 0.63) performed better on average than models built with features obtained from single timepoints (min: 0.55). Patient-level features (min: 0.56) were not sufficient in all testing scenarios as compared to intrapatient heterogeneity features (min: 0.63).

Conclusions: The candidate imaging biomarkers from previous ^{18}F-NaF PET/CT imaging studies of mPC patients did not generalize across all datasets. Incorporating response and heterogeneity features with single-timepoint and patient-level features resulted in RSF prediction models which were generalizable across all datasets. Use of such models hold promise for improving outcome prediction in mPC patients.