

**MACHINE LEARNING SIGNIFICANTLY IMPROVES
NEOANTIGEN-HLA PREDICTIONS UTILIZING > 26,000
DATA POINTS FROM THE PACTIMMUNETM DATABASE**

Vinnu Bhardwaj*, Amin Momin, Jonathan Johnston, Elizabeth Speltz, Tyler Borrman, Stefanie Mandl, Olivier Dalmás, Zheng Pan, Ashish Kheterpal, Eric Stawiski. *PACT Pharma, Inc, South San Francisco, CA, United States*

Background PACT Pharma has developed a state-of-the-art approach to validate predicted neoepitopes (neoEs) and their cognate T cell receptors (neoTCRs) by capturing neoepitope-specific T cells from peripheral blood. This neoTCR discovery and validation process is being applied in clinical trial (NCT03970382) evaluating personalized neoTCR-T cell therapy to treat patients across eight solid tumor types. Extensive pre-, on- and post-treatment data related to this trial has been accumulated in the PACTImmune Database (PIDB) which represents a growing data asset for patient-specific tumor immunogenicity in solid tumors. Here we present a specific use case of applying machine learning (ML) to significantly improve neoE-HLA predictions and further model anticipated improvements of TCR capture as a direct consequence.

Methods PACT has developed capabilities for high-throughput manufacturing of single polypeptide (comPACT protein) which consists of the predicted neoE peptide together with Beta-2-Microglobulin and the HLA heavy chain. comPACT molecules are considered successfully produced when protein yields reach concentrations >1 μ M. Data used for this study consisted of >26000 neoE-HLA predictions for 62 different HLA alleles. We applied ML to learn patterns that are predictive of neoE-HLAs that can be successfully produced as comPACTs, using scikit-learn and XGBoost. Data was first split into training and testing data. Models were trained on training data and model hyperparameters were tuned using 5-fold cross validation (5xCV). The performance of the models during 5xCV and on test data was measured using the area under the receiver operating characteristic curve (AUC). We additionally performed experimental prospective validation of the models. To do this, 603 neoE-HLAs (from 7 previously unseen cancer samples) were selected for comPACT production using netMHCpan4.1 and the newly trained models.

Results The mean AUC for the 5xCV of the selected models ranged from 0.75 to 0.86 depending upon the HLA allele (SD <0.05 for every model). The AUC on the test data ranged from 0.75 to 0.92 (median = 0.85). Prospective validation resulted on average in a 22% higher success rate (range 11%–39%) using the new models as compared to the netMHCpan4.1 predictions. This is expected to result in increased capture of neoepitope-specific CD8+ T cells as the PIDB indicates that 3.2% of the successful comPACTs result in validated neoTCRs.

Conclusions PIDB based ML predictions of neoE-HLAs led to a significant increase in TCR-capturing comPACT success rates. Because of this work, it is predicted both neoE-specific CD8+ T cell capture and actionable neoTCR options will increase per patient.

<http://dx.doi.org/10.1136/jitc-2021-SITC2021.820>