1300 PREDICTION OF EFFICACY AND TOXICITIES OF IMMUNE CHECKPOINT INHIBITORS USING REAL-WORLD PATIENT DATA

[1]Levente Lippenszky*, [2]Kathleen Mittendorf, [1]Zoltan Kiss, [3]Pablo Napan-Molina, [1]Eszter Csernai, [2]Michele LeNoue-Newton, [2]Christine Micheel, [4]Jan Wolber, [2]Travis Osterman. [1]*Science and Technology Organization – Artificial Intelligence and Machine Learning, GE HealthCare, Budapest, Budapest, Hungary*; [2]*Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center, Nashville, TN, USA*; [3]*Science and Technology Organization – Artificial Intelligence and Machine Learning, GE HealthCare, San Ramon, CA, USA*; [4]*Pharmaceutical Diagnostics, GE HealthCare, Chalfont St Giles, Buckinghamshire, UK*

Background Immune checkpoint inhibitors (ICI) have improved outcomes in several tumor types allowing subgroups of patients to have longer, higher quality lives. However, potential life-threatening immunotoxicities can arise in susceptible patients. Identifying patients at high risk of immunotoxicities alongside those responding well can help patients understand risk-benefit profile of the treatment, improve clinical trial cohort selection, and inform therapy selection in clinical settings. Herein, we introduce a machine learning (ML) framework that can accurately predict common immunotoxicities – hepatitis, colitis, and pneumonitis – alongside efficacy utilizing routinely collected Electronic Health Records (EHR) data.

Methods Our models rely on real-world EHR data of over 2,200 ICI-treated patients from Vanderbilt University Medical Center obtained prior to December 31, 2018. During the design of the predictive models, we set the prediction time point as the ICI initiation date for each patient. 1-year prediction time window was applied to create binary labels for the four prediction outcomes. Pneumonitis and colitis episodes were manually curated to establish the labels. The hepatitis label was defined to be 1 if any of the four liver enzymes exceeded three times the upper limit of normal. Overall survival served as a surrogate for efficacy. Structured data and clinician notes prior to ICI initiation were utilized to create features for the models. Feature engineering involved aggregating laboratory measurements over 60 and 120-day time windows. 1-year window was applied for other data types including ICD-10 codes, procedures, medication, and smoking history. In model development, patients were randomly partitioned into training (80%) and test (20%) sets for each outcome. An experiment involved a baseline and an alternative model, where the latter was selected if it demonstrated statistically significant superior performance based on outer loop results from a nested cross-validation process on the training set.

Results A random forest classifier was developed for each outcome. (table 1) demonstrates performance results with 95% bootstrap confidence intervals on the test set. Overall, each model shows reasonably strong performance achieving an AUC between 0.72 and 0.76. (table 2) contains the features used in the models.

Conclusions To our knowledge, this is the first ML solution that can assess the risk-benefit profile of ICI for patients, based on their medical history. As the models rely on routinely collected EHR data, their applicability does not require any changes in clinical practice. We envisage utility both in pre-screening of eligible patients for clinical trials and as clinical decision support in routine patient management.

Ethics Approval The Vanderbilt University Medical Center Health Sciences #3 institutional review board approved this study, tracked as #211814. The IRB determined the study poses minimal risk to participants, and a waiver of consent was granted.

Abstract 1300 Table 1 Model performance results

Prediction outcome	AUC	Sensitivity	PPV
Pneumonitis 1y	0.739 (0.638 -0.823)	0.711 (0.561-0.845)	0.193 (0.129-0.261)
Hepatitis 1y	0.729 (0.655-0.804)	0.455 (0.333-0.574)	0.395 (0.284-0.508)
Colitis 1y	0.755 (0.638-0.856)	0.750 (0.581-0.909)	0.134 (0.086-0.182)
Overall survival 1y	0.752 (0.706-0.796)	0.817 (0.764-0.865)	0.664 (0.609-0.719)

Abstract 1300 Table 2 Features of the models

	Data	Data type	Aggregation	Time window
Pneumonitis 1y	C34, C78, R91, J, R05, R06, R07, R09	ICD-10 condition codes	Relative frequency	1 year
	C34	ICD-10 condition codes	Relative frequency	
	C78	ICD-10 condition codes	Relative frequency	
	J44 indicator	ICD-10 condition codes	One-hot encoding	
	Smoking indicator	Curation	One-hot encoding	
	Interaction between smoking indicator and (C34 or C78) indicator	Curation and ICD-10 condition codes	One-hot encoding	
	Oxygen saturation in blood	Laboratory measurements	Time-weighted average	120 days
	Body mass index	Body measurements	Time-weighted average	
Hepatitis 1y	Aspartate aminotransferase	Laboratory measurements	Min, max, last, time-weighted average	60 days
	Alanine transaminase			
	Alkaline phosphatase			
	Bilirubin			
Colitis 1y	K50, K51, K52, K57, K58	ICD-10 condition codes	Relative frequency	1 year
	Chemotherapy indicator	Procedures	One-hot encoding	
	C43 indicator	ICD-10 condition codes	One-hot encoding	
	CTLA-4 indicator	Drugs	One-hot encoding	
	PD-1 indicator	Drugs	One-hot encoding	
	Number of ICI drugs	Drugs	Count	
	Albumin	Laboratory measurements	Relative frequency below and above normal range	120 days
	Hemoglobin			
	Red blood cell			
	Absolute lymphocytes count			
	White blood cell			
OS 1y	Chemotherapy indicator	Procedures	One-hot encoding	1 year
	Albumin	Laboratory measurements	Last, time-weighted average	120 days
	Hemoglobin			
	Neutrophils			
	Absolute lymphocytes count			
	Alkaline phosphatase			
	Alanine transaminase			
	Packed cell volume			
	Oxygen saturation in blood			
	C78	ICD-10 condition codes	Relative frequency	1 year
	Age	Demographics	Value in years	At first ICI

http://dx.doi.org/10.1136/jitc-2023-SITC2023.1300