FPN 501P A Prognostic Model Of All-cause Mortality At 30 Days In Patients With Cancer And COVID-19

Susan Halabi, PhD[1]*, Bin Luo, PhD[1], Hannah Dzimitrowicz[1], Clara Hwang, MD[2], Trisha M. Wise-Draper, MD, PhD[3], Chris Labaki, MD[4], Rana R. McKay, MD[5], Erika Ruíz-García, MD, MCs[6], Claudia Rangel-Escareño, PhD[7,8], Dimitrios Farmakiotis, MD[9], Elizabeth A. Griffiths, MD[10], Chinmay Jani, MD[11], Melissa K. Accordino, MD, MS[12], Christopher R. Friese, PhD, RN[13], Elizabeth Wulff-Burchfield, MD[14], Matthew Puc, MD[15], Peter Paul Yu, MD[16], Umit Topaloglu, PhD[17], Sanjay Mishra, PhD[18], Jeremy L. Warner, MD, MS[18,19]

[1]Duke Cancer Institute at Duke University Medical Center; [2]Henry Ford Cancer Institute, Henry Ford Hospital; [3]University of Cincinnati Cancer Center; [4]Dana-Farber Cancer Institute; [5]Moores Comprehensive Cancer Center at the University of California, San Diego; [6]Instituto Nacional de Cancerologia; [7]National Institute of Genomic Medicine, Mexico; [8]Tecnologico de Monterrey; [9]Brown University and Lifespan Cancer Institute; [10]Roswell Park Comprehensive Cancer Center; [11]Mount Auburn Hospital; [12]Herbert Irving Comprehensive Cancer Center at Columbia University; [13]University of Michigan Rogel Cancer Center; [14]The University of Kansas Cancer Center; [15]Virtua Health; [16]Hartford HealthCare Cancer Institute; [17]Wake Forest Baptist Comprehensive Cancer Center; [18]Vanderbilt-Ingram Cancer Center, Vanderbilt University Medical Center; [19] Brown University.

Background and Objectives

- Patients with cancer are at higher risk of dying of COVID-19
- Known risk factors for 30-days all-cause mortality (ACM-30) in patients with cancer are older age, gender, smoking status, performance status, obesity, and co-morbidities
- **Objectives:** 1) Identify informative common clinical and laboratory parameters predictive of a higher risk of 30-days ACM; 2) Build and validate a machine learning model based on clinical and laboratory values to estimate individual survival probability of ACM at 30 days

Methods

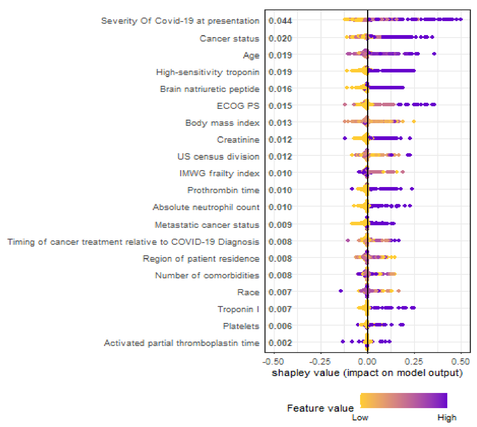
- COVID-19 and Cancer Consortium (CCC19) international registry study
- 12,668 patients were enrolled between March 17, 2020 and December 31, 2021
- 891 patients were excluded due to quality score>4 (poor quality data)
- Data collected & managed using REDCap at Vanderbilt University Medical Center

Data Analysis

- ACM-30 days defined death from any cause within 30 days of COVID19 diagnosis
- Pre-specified variables : age, sex, race, smoking status, ECOG performance status, timing of cancer treatment relative to COVID19 diagnosis, severity of COVID19, type of cancer, and other laboratory measurements
- Imputed missing variables using random forest proximity
- Randomly split data into training and testing sets allocation ratio 2:1 stratifying on US census division
- Two-step random forest utilized to model ACM-30: top twenty variables first selected based on feature importance given by random forest; model was refitted using the selected variables on the training set
- Area under the curve (AUC) computed measure of predictive accuracy with out-of-bag (OOB) and testing prediction.
- Optimal predicted probability cutoff ACM-30 selected maximum Youden's Index
- Sensitivity and specificity computed under the optimal probability cutoff
- Monte Carol cross-validation: repeated random splitting and model fitting 100 times and obtained variable selection frequency and distribution of the AUCs

Variable Contribution

Shapley plot showing the contributions of the twenty selected variables on the prediction of the final model

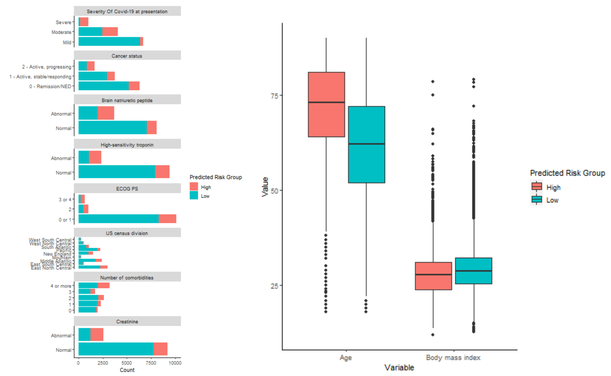


Results

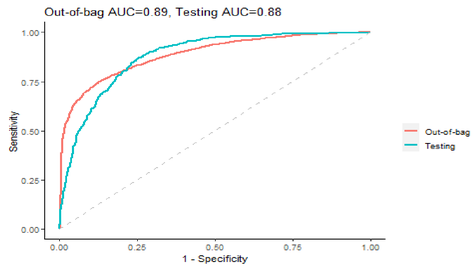
Baseline Characteristics

ACM-30	No (N=10,343)	Yes (N=1,423)	Overall (N=11,766)
Age			
Median (Q1, Q3)	64.0 (54.0, 73.0)	73.0 (64.0, 81.0)	65.0 (55.0, 74.0)
Missing (%)	0.2	0.2	0.2
Body mass index			
Median (Q1, Q3)	28.3 (24.6, 32.9)	26.9 (23.0, 32.0)	28.1 (24.4, 32.9)
Missing (%)	12.6	17.1	13.2
Sex (%)			
Female	54.2	44.1	53.0
Male	45.7	55.8	46.9
Missing	0.1	0.1	0.1
Race (%)			
Hispanic	17.8	15.8	17.6
Non-Hispanic Black	17.1	21.2	17.6
Non-Hispanic White	52.3	52.1	52.3
Other	11.1	9.8	10.9
Missing	1.7	1.2	1.7
Smoking status (%)			
Current	6.1	5.3	6.0
Former	36.1	47.9	37.5
Never	54.3	41.2	52.7
Missing	3.5	5.6	3.7
ECOG Performance Status (%)			
0 or 1	60.8	37.3	57.9
2	7.4	19.2	8.8
3 or 4	3.6	17.7	5.3
Missing	28.3	25.8	28.0
Cancer status (%)			
0 - Remission/NED	48.8	32.1	46.8
1 - Active, stable/responding	29.6	24.0	28.9
2 - Active, progressing	11.6	30.8	13.9
Missing	10.0	13.1	10.4
Severity Of Covid-19 (%)			
Mild	62.9	11.4	56.7
Moderate	31.8	53.6	34.4
Severe	5.1	34.7	8.7
Missing	0.2	0.3	0.2
US Census Division (%)			
East North Central	22.8	20.8	22.6
East South Central	4.7	4.6	4.7
Middle Atlantic	19.5	23.5	20.0
Mountain	2.3	1.0	2.1
New England	12.1	16.4	12.6
Pacific	14.9	8.9	14.2
South Atlantic	8.9	8.2	8.8
West North Central	4.6	3.2	4.4
West South Central	2.1	2.5	2.2
Missing	8.0	11.0	8.3
Timing of cancer treatment relative to COVID-19 Diagnosis (%)			
More than 3 months	44.2	36.4	43.3
0-4 weeks	35.3	37.0	35.5
1-3 months	8.0	12.4	8.5
Never/after COVID-19 diagnosis	7.9	8.9	8.1
Missing	4.6	5.3	4.7

Predicted Risk Group by Selected Variables



Prediction Accuracy



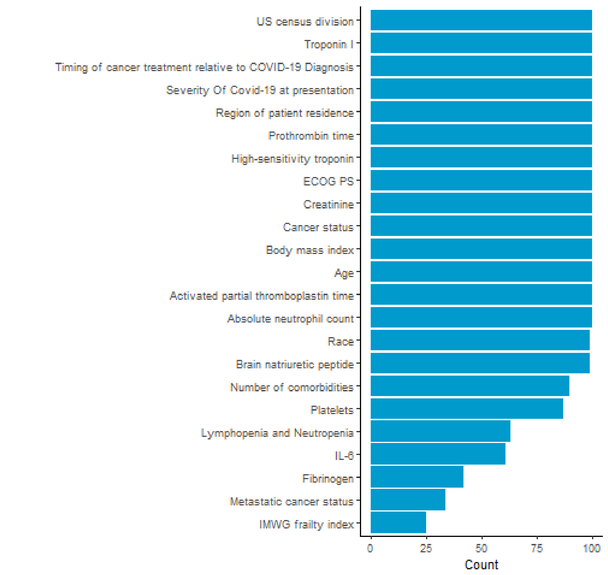
Observed vs. predicted ACM at 30 days on Testing set

	ACM-30 Days	
Predicted ACM-30	Yes	No
Yes	415	865
No	63	2584

- Optimal probability cutoff maximizing Youden's index=0.127
- The sensitivity and specificity on the testing set are 0.87 and 0.75

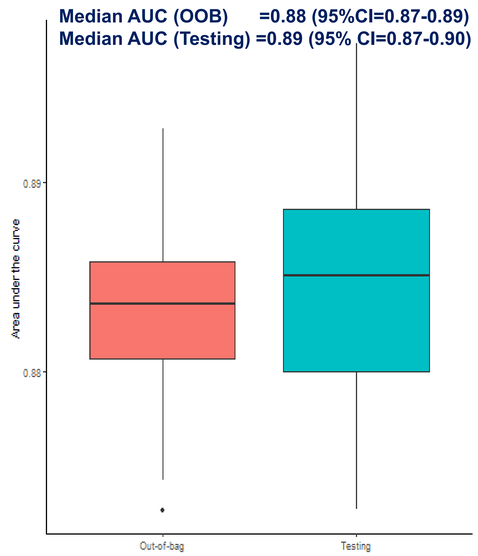
Monte Carlo Cross Validation

Frequencies of variable selection over 100 data splits



Distribution of AUCs over 100 data splits

Median AUC (OOB) =0.88 (95%CI=0.87-0.89)
Median AUC (Testing) =0.89 (95% CI=0.87-0.90)



Conclusion

- Top predictors: Severity of COVID19, cancer status and age
- ACM-30 days model includes readily available clinical and laboratory variables
- Model can be used to estimate individual survival probability within 30-days for COVID-19
- Model can be used to select or classify patients with cancer and COVID-19 into risk groups based on validated cut points, for treatment selection, prophylaxis prioritization, and/or enrollment in clinical trials.
- Future work includes validation from external dataset of patients with COVID-19 and cancer

Acknowledgement

We thank the CCC19 Steering Committee and Investigators for their contribution to this register. Funding: CCC19 supported by NCI P30 CA068484 to Vanderbilt University.