Leveraging Precision Medicine: COVID19 in the ER

Tom Velez PhD Esq, Tony Wang PhD, Brian Garibaldi MD, Eric Singman MD PhD & Ioannis Koutroulis MD PhD MBA

Background: There is significant heterogeneity in disease progression among COVID19 patients reaching the ER. This is attributed to a complex interplay between virus and host immune response that can unpredictably and rapidly lead to potentially fatal "hyperinflammation." Early identification of patients at risk of hyperinflammation can guide decisions to step up to ICU care or not. This saves lives and healthcare dollars.

Objective: The primary objective of this study is to use machine learning to reproducibly identify specific risk-stratifying clinical phenotypes across hospitalized COVID19 patients and compare treatment response characteristics and outcomes. A secondary objective is to derive a predictive phenotype classification model using only routinely available early encounter (<6 hr) data that may be useful in informing optimal COVID19 bedside clinical management.

Methods: This is a retrospective analysis employing the **CROWN COVID Registry** (de-identified EPIC records of adults that were admitted to the Johns Hopkins hospitals for COVID19) in the 2020-2021 timeframe. Phenotypes were identified by clustering 38 routine clinical observations recorded during first 6 hours of ED/inpatient care. To examine the reproducibility/validity of derived phenotypes, patient data was randomly divided into two cohorts; clustering analysis was performed independently on each cohort. A predictive phenotype classifier using the Gradient Boosting Machine (GBM) method was derived using these data. .

Results: Two phenotypes (P1 and P2) were identified in patients admitted for COVID19 in both the training and validation cohorts with similar distributions of features, correlations with biomarkers, treatments, comorbidities, and outcomes. In both training and validation cohorts, P2 patients were **older, had elevated markers of inflammation and were at an increased risk of requiring ICU-level care, developing sepsis, and mortality**. The GBM phenotype predictive model yielded an area under the curve (AUC) of 0.89 and a positive predictive value (PPV) of 0.83.

Conclusions: Using machine learning clustering of only routine early-admission data, we identified and validated two COVID19 phenotypes with distinct treatment/response characteristics consistent with similar two-phenotype models derived in other hospitalized COVID19 populations, supporting the reliability and generalizability of these findings. A phenotype predictive model based on early encounter data may be clinically useful for timely bedside risk stratification and treatment personalization.

Ramifications:
Delayed ICU admission when required is associated with significantly higher mortality. What of unnecessarily premature ICU admission?
15-20% of COVID19 patients require ICU-care & ICU costs are estimated to be 3-7x greater than general ward care.
A one-day substitution of general ward care for ICU care results in an estimated cost reduction of at least $1200/patient.
There were 472345 COVID- associated hospital admissions from 2020-2023, with an estimated 71000 ICU admissions.
If even 5% could safely delay ICU admission by 1 day, the cost savings would be at least **$4.26M. Predicting need for ICU care within 6 hours of ER admission saves lives and money.**

Reference: Velez T, Wang T, Garibaldi B, Singman E, Koutroulis I. Identification and Prediction of Clinical Phenotypes in Hospitalized COVID-19 Patients: A Step towards Precision Medicine. JMIR Form Res. 2023 Aug 24. doi: 10.2196/46807. Epub ahead of print. PMID: 37642512.

Features used for clustering

Vitals	Min	SpO2, SpO2/FiO2, Systolic BP, Pulse pressure
	Max	Pulse, Respiratory Rate, Temperature
Labs	Min	Albumin, Ca^{++}, CO_2, Gamma gap, HCT, Hgb, Lymphs, MCH, MCV, Monocytes, Platelets, Protein, K^+, RBC, RDW, Na^+
	Max	AST, ALT, Anion Gap, Bilirubin, BUN/Cr, CRP, Glucose, MPV, Neutrophils, Neutro/Lymph Ratio, Pltlt/Lymph Ration, WBC

Outcomes/Treatments by Phenotype

Characteristics	Training cohort			Validation cohort		
	Phenotype 1	Phenotype 2	P	Phenotype 1	Phenotype 2	P
N	1284	898		1258	939	
Sepsis, n (%)	257 (20.0)	306 (34.1)	<0.001	273 (21.7)	315 (33.5)	<0.001
Ventilation, n (%)	53 (4.1)	97 (10.8)	<0.001	63 (5.0)	105 (11.2)	<0.001
IV pressor, n (%)	41 (3.2)	98 (10.9)	<0.001	49 (3.9)	94 (10.0)	<0.001
Hi Flo O_2, n (%)	99 (7.7)	155 (17.3)	<0.001	124 (9.9)	139 (14.8)	<0.001
Renal Replace, n (%)	4 (0.3)	20 (2.2)	<0.001	9(0.7)	24 (2.6)	<0.001
Dialysis, n (%)	3 (0.2)	84 (9.4)	<0.001	12(1.0)	74 (7.9)	<0.001
Death, n (%)	25 (1.9)	150 (16.7)	<0.001	39 (3.1)	160 (17.0)	<0.001

Prediction model performance

Metric	Estimate	95% CI
AUC	0.890	0.887-893
Sensitivity	0.846	0.822-0.873
Specificity	0.851	0.828-0.876
PPV	0.834	0.807-0.865
NPV	0.861	0.845-0.879

Disclosure: Dr. Velez is CEO of Vivace Health Solutions, software developer for this machine learning effort.