PB2477

Expanded studies of a methylation-based COVID-19 classification model to predict severity of disease and its ability to differentiate from other respiratory viruses

Brett R. Peterson[1], Wenyu Zhou[1], Genelle F. Harrison[1], Meher Preethi Boorgula[2], Monica Campbell[2], Sameer Chavan[2], Bret Barnes[3], Rishi Porecha[3], Rasika A. Mathias[4], Ivana Yang[2], Christopher Gignoux[2], Alem Taye[3], Andrew Monte[2], Kathleen C. Barnes[1,2,4]

[1]Tempus Labs, Chicago, IL, [2]University of Colorado, Denver, CO, [3]Illumina Inc. La Jolla, CA, [4]Johns Hopkins University, Baltimore, MD

INTRODUCTION

- Recently we customized Illumina's Infinium MethylationEPIC array (Epic+) to enhance immune response detection and profiled blood samples from patients with/without SARS-CoV-2 infection to (Fig. 1; Konigsberg et al 2021 Comm Med).
- We designed a machine learning (ML) platform that leveraged methylation risk scoring (MRS) and demonstrated that host methylation predicts SARS-CoV-2 infection and clinical outcome.
- Previously we compared SARS-CoV-2+ patients to 65 patients with other respiratory viral infections to determine the specificity of the methylation signature to SARS-CoV-2. In the current study, we modified the model to improve prediction of both SARS-CoV-2 and other viral infections.
- We further hypothesized that a larger dataset might increase the sensitivity of our model overall.

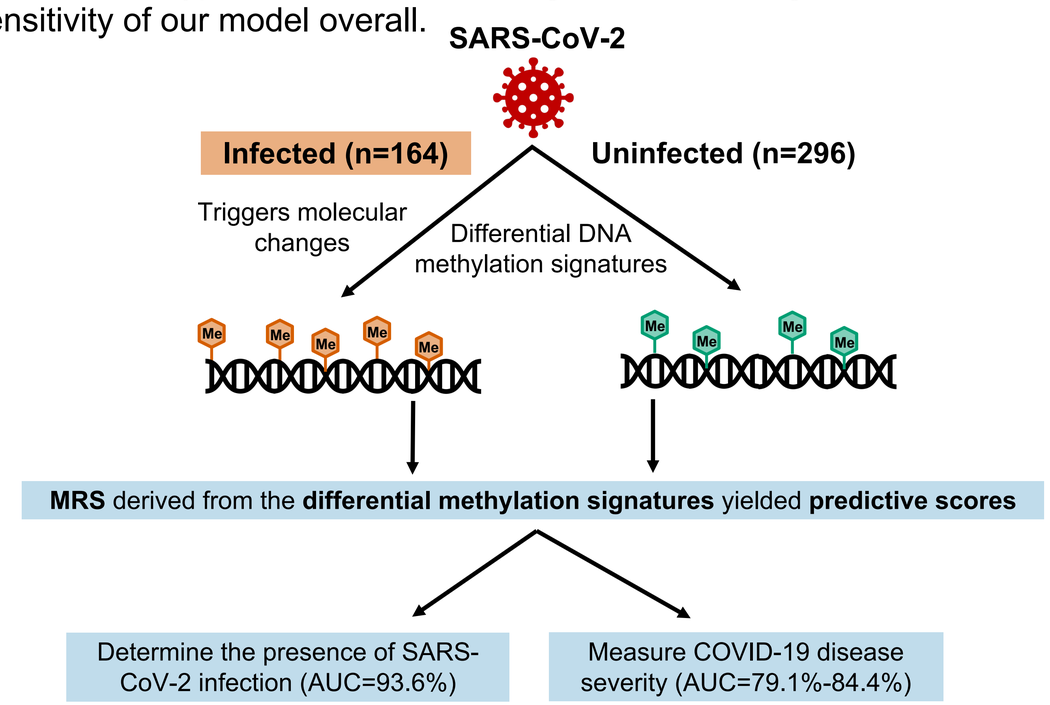


Figure 1. Phase 1 development of an MRS for SARS-CoV-2 status and severity

METHODS

- **Phase 2 Data Collection:** Genome-wide DNA methylation was measured in an additional 135 SARS-CoV-2 cases and 88 controls using the customized Epic+ array (**Table 1**).
- Phase 2 includes 66 SARS-CoV-2$^-$ patients with other respiratory viruses, covering 8 types of distinct viral families.
- Phenotype data were extracted from the electronic health record (EHR) for disease progression and severity ((hospitalization, ICU admittance, ventilator use, and patient death) and 'other respiratory virus infection'.

Trait	Total (N=741)	COVID -ve (N=376)	COVID +ve (N=299)	Other Viral Infections (N=66)
Mean Age, years (SD)	52.8 (18.3)	54.3 (18.7)	50.4 (17.7)	54.7 (17.5)
Female, Sex, n(%)	359 (48.4%)	187 (49.7%)	136 (45.5%)	36 (54.5%)
Ancestry, n(%)				
Native American or Alaskan Native	4 (0.5%)	2 (0.5%)	2 (0.7%)	0 (0%)
Asian	24 (3.2%)	7 (1.9%)	17 (5.7%)	0 (0%)
Black or AA	143 (19.3%)	66 (17.6%)	60 (20.1%)	17 (25.8%)
Native Hawaiian or Pacific Islander	9 (1.2%)	5(1.3%)	4 (1.3%)	0 (0%)
White	361 (48.7%)	219 (58.2%)	103 (34.4%)	39 (59.1%)
Other	178 (24.0%)	66 (17.5%)	103 (34.4%)	9 (13.6%)
Unknown or not reported	22 (3.0%)	11 (2.9%)	10 (3.3%)	1 (1.5%)
Hispanic Latino, n(%)	222 (29.9%)	81 (21.5%)	129 (43.1%)	12 (18.1%)
Emergency Department Disposition, n (%)				
Discharged	106 (14.3%)	48 (12.7%)	58 (19.4%)	0 (0%)
Floor admission	104 (14.0%)	28 (7.4%)	75 (25.0%)	1 (1.5%)
Number ever admitted to the ICU during the hospital stay	128 (17.3%)	73 (19.4%)	45 (15.0%)	10 (15.1%)
Number ever requiring ventilator during hospital stay	74 (10.0%)	44 (11.7%)	28 (9.4%)	2 (3.0%)

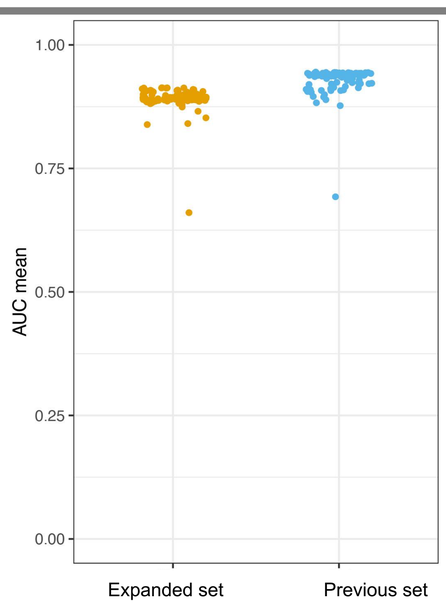
Table 1. Demographics for the expanded (Phase 2) cohort.

SIGNIFICANCE

- The **COVID-19-specific epigenetic signature** from peripheral blood driven by expression/activation of key immune-related pathways is **related to infection status, disease severity, and clinical deterioration in our expanded dataset**.
- Compared to the previously reported sparse regression based MRS (prev_model), the new MRS model trained in this expanded cohort (expd_model) did not significantly improve AUCs for case-vs-control status, hospitalization, ICU admission, and progression to death. We did, however, **observe higher predictive scores for other respiratory viruses with the larger dataset**.

RESULTS

Figure 1. The predictive model is unchanged when additional COVID-19 cases and controls are added. Here we show the distribution of the AUC mean during the cross-validation training either in previous set (70% for training, case=114, control=207) or expanded set (80% for training, case=237, control=305), excluded other infections in training. Previously we observed AUC=0.93(+/- 0.019); in the expanded set we observe AUC=0.92 (+/- 0.02) with no significant difference observed (t-test p = 0.43).



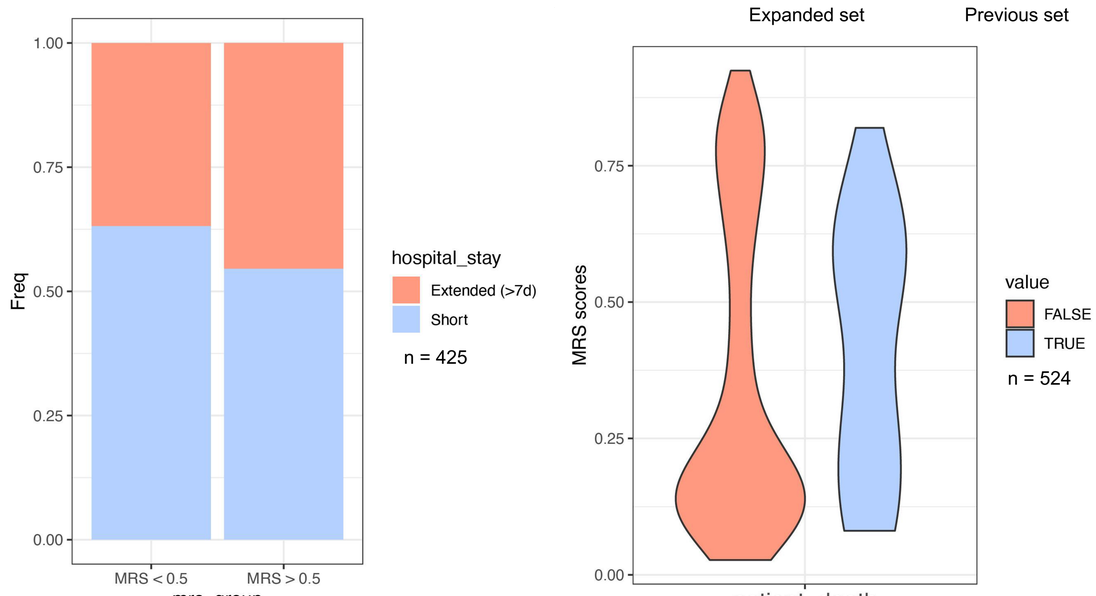
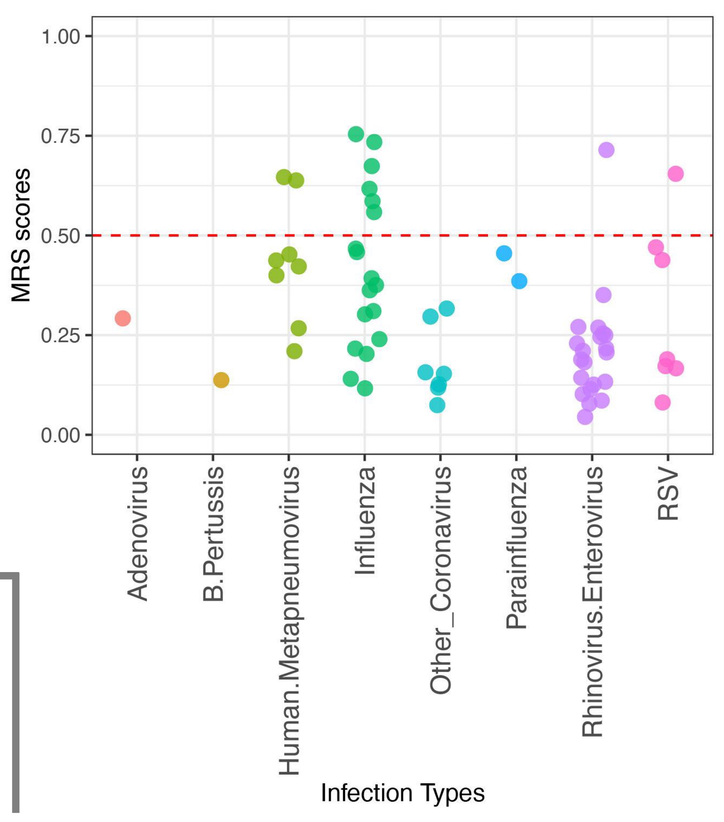


Figure 2. MRS scores predict COVID-19 severity. Here we show that **(left)** patients with higher MRS scores have a higher proportion of extended hospital stay (>7 days); and **(right)** patients with fatal outcomes have higher MRS score than patients who recovered.

Figure 3. The COVID-19 model shows various levels of specificity for other respiratory viral infections. Here we show distribution of MRSs in other infections among COVID-19 negative controls.



Acknowledgements: This work was funded by NIH/NIAID U19 AI117673 & funding from the University of Colorado Anschutz Medical Campus Dept of Medicine. We thank Amrita A. Iyer, Ph.D. for poster preparation and review, and the Scientific Communications and Design teams at Tempus for data visualization guidelines & poster review.

Correspondence: kathleen.barnes@tempus.com / kathleen.barnes@cuanschutz.edu