Genetic and inflammatory signatures associated with worse prognosis in hospitalized patients with severe SARS-CoV-2 infection and diabetes.

Marshall Yuan, Andrew Wassef, Davit Sargsyan, Vahe Nersisyan, Javier Cabrera, Ronald G, Nahass, Ah-Ng Kong, Luigi Brunetti

Robert Wood Johnson Medical School, Piscataway, NJ, USA

IDCare, Hillsborough, NJ, USA

Ernest Mario School of Pharmacy, Department of Pharmaceutics, Piscataway, NJ, USA

Ernest Mario School of Pharmacy, Department of Pharmacy Practice, Piscataway, NJ, USA

Robert Wood Johnson University Hospital, Somerville, NJ, USA

Keywords

Word Count

Table of Contents

[Abstract 3](#_Toc159255866)

[Introduction 4](#_Toc159255867)

[Methods 5](#_Toc159255868)

[Data source and sample collection 5](#_Toc159255869)

[Data extraction and collection 5](#_Toc159255870)

[RNA-Sequencing 5](#_Toc159255871)

[Cytokine and chemokine multiplex assay 5](#_Toc159255872)

[ACE2 and DPP-IV ELISA 6](#_Toc159255873)

[Clinical Outcomes 6](#_Toc159255874)

[Statistical Analysis 6](#_Toc159255875)

[Results 6](#_Toc159255876)

[Clinical outcomes 7](#_Toc159255877)

[Inflammatory signature analysis 8](#_Toc159255878)

[DPP-IV and ACE2 signature analysis 8](#_Toc159255879)

[Differential Gene Expression Analyses 8](#_Toc159255880)

[Discussion 13](#_Toc159255881)

[References 13](#_Toc159255882)

# Abstract

# Introduction

Coronavirus Disease of 2019 (COVID-19), caused by the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), was announced as a pandemic by the WHO at the beginning of 2020 due to its rapid communicability and disease severity. By October 2022, COVID-19 had caused over 6.5 million deaths. Primarily a condition that affects the respiratory system, it presents in patients with a wide range of symptoms, ranging from asymptomatic and mild to severe. In the most critical cases, patients may require ICU care and mechanical intubation, among other intensive interventions. A variety of risk factors are suggested to increase the risk for severe illness, including 65 years of age and older, hypertension, smoking, and diabetes.

Multiple meta-analyses of the clinical correlation between diabetes and SARS-CoV-2 have demonstrated that patients with diabetes are at higher risk for severe disease and mortality, reporting odds ratio as high as OR = 2.75 (95% CI: 2.09-3.62; p < 0.01) for severe disease. Diabetes has been previously implicated in other infectious conditions, including being associated with over a four-fold risk of ICU admission in patients with the Influenza A infection of 2009 (H1N1). Furthermore, diabetes has been observed to be associated with critical illness and identified as an independent risk factor for 90-day mortality in patients with Middle East respiratory syndrome coronavirus (MERS-CoV). Other studies further corroborate a bi-directional link between diabetes and COVID-19, including cases and systematic reviews that found a higher incidence rate of new-onset diabetes and hyperglycemia in patients previously infected by COVID-19. Despite the substantial data that supports diabetes as a risk factor for diabetes, the mechanism that mediates this risk is largely unknown.

Although poorly elucidated, the mechanism of disease severity in diabetes mellitus patients may be connected to angiotensin-converting enzyme 2 (ACE2) and cytokine/chemokine gene expression. Severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) uses the ACE2 receptor to enter host cells. Upon entry, there is a downregulation of surface ACE2 expression. Circulating angiotensin 2 (Ang-II) are elevated in COVID-19 patients compared to healthy controls providing evidence of renin-angiotensin system (RAS) imbalance in the disease. Increases in Ang-II leads to increases in a disintegrin and metalloproteinase 17 (ADAM17) activity and subsequent release of tumor necrosis factor α (TNF- α) and other inflammatory cytokines. Nuclear factor erythroid 2–related factor 2 (NRF2) and NRF2 related genes are regulators of cellular redox balance and are involved in the release of inflammatory cytokines and chemokines secondary to stress. NRF2 activation downregulated a variety of cytokines reported to be elevated in COVID-19 suggesting reduced NRF2 activity as a contributor to the “cytokine storm” seen in COVID-19.

Chemokines are an important secretory protein responsible for immune signaling and have been implicated in a variety of lung pathologies. For example, CCL2 [chemokine (C-C motif) ligand 2; monocyte chemoattractant protein-1, (MCP-1)] and its receptor CCR2 are involved in monocyte/macrophage migration, Th2 cell polarization, and the production of TGF-β and procollagen in fibroblast cells. This chemokine is associated with acute respiratory distress syndrome and pulmonary fibrosis7 – both observed in COVID-19. CCL2 elevation has also been found to be associated with severe SARS-CoV. A variety of chemokines have been reported to be elevated in COVID-19 infection, but there has not been an evaluation of differential patterns of expression in individuals with and without diabetes.

The purpose of this study was to evaluate biomarker and gene expression patterns in individuals hospitalized with diabetes mellitus infected with SARS-CoV-2. In addition, the relationship between these patterns and disease severity was examined.

# Methods

Data source and sample collection

We performed a single-center, IRB-approved, cohort study using data from electronic health records and left over clinical specimens at a large community medical center. All subjects 18 years of age or older with remnant clinical blood specimens within 48 hours of hospital admission were eligible for inclusion. Pregnant patients and those discharged directly from the Emergency Department were excluded. For each patient, an aliquot of whole blood sample collected in an EDTA tube (500 uL per sample) was mixed with 1.3 mL of RNALater. The remaining whole blood was centrifuged at 3000 x *G* for 10 minutes, and plasma was drawn off. All samples were stored at -80°C until analysis.

## Data extraction and collection

All data were extracted from the electronic health record (EPIC Systems; Wisconsin, USA). Patient age, sex, race/ethnicity, comorbidities, vaccination status, concomitant anti-hyperglycemic medications, COVID-related treatment interventions, and other basic relevant clinical laboratory data were extracted from the records. Patient comorbidities were identified using the International Classification of Diseases, tenth revision, clinical modification (ICD-10-CM) codes. The overall comorbidity status of patients was defined by the scoring of the Charlson-Deyo comorbidity Index (CCI).

## RNA-Sequencing

RNA sequencing was performed by Singulomics. Briefly, Messenger RNA was purified from total RNA using poly-T oligo-attached magnetic beads. After fragmentation, the first strand cDNA was synthesized using random hexamer primers, followed by the second strand cDNA synthesis using either dUTP for directional library or dTTP for non-directional library. For the non-directional library, it was ready after end repair, A-tailing, adapter ligation, size selection, amplification, and purification

## Cytokine and chemokine multiplex assay

Plasma cytokine and chemokine concentrations were measured using ProcartaPlexTM Human Cytokine Storm 21-Plex (Invitrogen; EPX210-15850-901) on the Luminex platform. Briefly, 25 µL of plasma and internal controls were plated on a 96-well plate, incubated with magnetic beads, and washed before the addition of 25 μL of detection antibody. The plate was then incubated for 30 min followed by the addition of 50 μL of Streptavidin-PE to each well. The concentration of analytes was reported as pg/mL.

## ACE2 and DPP-IV ELISA

Circulating ACE2 was measured by sandwich ELISA (Invitrogen; EH489RB). Briefly, 100 μL of standards and 100 μL of diluted plasma samples were plated on a 96-well plate. After a series of washes, 100 μL of biotin was added to each well, followed by a 1-hour incubation period at room temperature. The solution was discarded, and the plate was washed, followed by the addition of 100 μL of streptavidin-HRP. The plate was incubated for 45 minutes. After the solution was discarded and the plate washed, 100 μL of TMB substrate was added and incubated for 30 minutes. Once stop solution was added, the plate was read at an absorbance of 450 nm. The concentration of plasma ACE2 was reported as ng/mL.

Similarly, DPP-IV was measured using a sandwich ELISA (InvitrogenTM; EHDPP4). The concentration of plasma DPP-IV was reported as pg/mL

## Clinical Outcomes

Patients were stratified into those with COVID-19 and those without COVID-19 as well as those with diabetes and those without diabetes. The primary endpoint was differences in inflammatory mediator expression profile between COVID-19 patients with and without diabetes. Secondary endpoints include differences in inflammatory mediator expression profile between diabetes patients with and without COVID-19 and occurrence of severe illness (defined by mechanical ventilation, ICU admission, or mortality).

## Statistical Analysis

Statistical analysis and visualization were conducted using *R* 4.3.1 software (REFERENCE: R: A Language and Environment for Statistical Computing. R Core Team, R Foundation for Statistical Computing. Vienna, Austria, 2023. https://www.R-project.org/).

Patient demographics were assessed using descriptive statistics. Continuous data were represented by means and standard errors, while nominal data were presented as percentages. Differences in baseline characteristics were analyzed utilizing t-test or analysis of variance (ANOVA) for continuous data, and chi-squared or Fischer’s exact test for categorical data.

RNA-seq data was analyzed with DESeq2 R package (REFERENCE: Love, M.I., Huber, W., Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2 Genome Biology 15(12):550 (2014)) for differential gene expression analysis based on the negative binomial distribution. Pathway analysis was conducted with ReactomPA package (REFERENCE: Yu G, He Q (2016). “ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization.” Molecular BioSystems, 12(12), 477-479. doi:10.1039/C5MB00663E, <http://pubs.rsc.org/en/Content/ArticleLanding/2015/MB/C5MB00663E>).

# Results

The study included 182 hospitalized adult patients with an admitting diagnosis of COVID-19 (n=110) and control subjects admitted for any other reason (n=72). Remnant clinical blood samples were available for 91 patients (XX with COVID-19 and XX without). All blood samples were obtained within 48 hours of hospital presentation. Table 1 summarizes patients’ baseline characteristics, including demographics, vaccination information, comorbidities, medications, baseline lab work, and clinical scores, stratified by diabetes mellitus (DM) diagnosis

Out of 110 patients admitted with COVID-19, 12 (10.9%) died during hospitalization compared to 3 out of 72 non-COVID-19 patients (4.2%). Three (3) out of the 12 COVID-19 patients that died in hospital were admitted for shortness of breath (ICD10 R06.02), unspecified fever (ICD-10 R50.9) or fatigue (ICD-10 R53.83). The 3 non-COVID in-hospital deaths occurred in patients admitted for pneumonia (ICD-10 J18.9) or acute respiratory distress (ICD-10 R06.03).

90-day mortality (after discharge) was recorded for 5 non-COVID patients who were initially admitted for acute ischemic heart disease (ICD-10 I24.9), pneumonia, acute respiratory distress, or intestinal obstruction (ICD-10 K56.609). Additionally, 26 non-COVID patients were readmitted to hospitals within 90 days of discharge. No follow-up information was available for the COVID-19 patients at the time of this report.

## Clinical outcomes

There was no significant association between DM and COVID-19 patients’ in-hospital death rate, with 6 DM and 6 non-DM COVID-19 patients dying in hospital (Chi-square test p-value = 0.236). Similarly, obesity and BMI were not found to be significant factors associated with in-hospital death (p-values of 0.760 and >0.999, respectively). However, the odds of in-hospital death were 21.5 times higher (95% CI = 5.2 to 88.3, p-value< 0.001) for the COVID patients admitted to the critical care unit (ICU) compared to those who were not admitted to ICU. Specifically, 7 out of 13 ICU-admitted patients died in hospital compared to 5 deaths occurring in 97 non-ICU patients.

The odds ratio of in-hospital death for patients admitted with COVID-19 versus non-COVID-19 patients was not statistically significantly different from 1 (OR=2.82, 95%CI = 0.86 to 12.70, p-value = 0.119). However, after adjusting for ICU, the association of COVID-19 diagnoses with death became significant (OR = 6.79, 95%CI = 1.73 to 36.07, p-value = 0.012).

COVID-19 severity was measured on World Health Organization Original Scale for Clinical Improvement (WHO OSCI) scale. COVID-19 patients were grouped by the WHO OSCI into Moderate (score < 5) and severe (>=5 and <8) cohorts. WHO OSCI score of 8 signified deaths. At the admission, 77 out of 110 COVID patients had WHO OSCI score of 5 or higher. Notable, 5 out of the 13 COVID patient admitted to the CC unit had the score of 5, and another 8 score of 6. At the same time, 64 out of 97 non-CCU patients (66.0%) were scored 5 or above at the admission. Additionally, all 12 COVID-19 patients that died in hospital had WHO OSCI score of 5 or 6 at the admission, and their scores did not decrease until their death except for a single patient whose score decreased from 5 to 4 on Day 3, just before death (Table 2).

On average, COVID patients were admitted for a slightly shorter period compared to non-COVID patients (mean+/-SEM = 7.3+/-0.9 and 8.8+/-1.1 days respectively). The patients who died in hospital were hospitalized for longer times compared to those discharged alive (12.5+/-2.4 vs 7.5+/-0.7, respectively). The difference between COVID and non-COVID patients’ length of stay was even larger for those who were not discharged on the day of admission (i.e., stayed for more than one day), with LOS of 10.1+/-1.2 days for non-Covid patients discharged alive vs. 7.4+/-0.9 days for the COVID patients discharged alive. For the patients who died in hospital, the LOS were 11.0+/-3.8 and 12.8+/-2.9 for non-COVID vs COVID patients, respectively.

Patients’ comparison (medications, treatments administered, length of stay, vaccination status)

Outcomes (inflammatory mediator expression profile, statistical analysis)

## Inflammatory signature analysis

Proteins were measured using ELISA assay. In total, 21 protein concentrations were measured in 54 COVID and 68 non-COVID patients.

*(Need to compare cytokines and chemokines between COVID No COVID; COVID with and without DM AND; then between COVID with DM and between NO COVID and DM)*

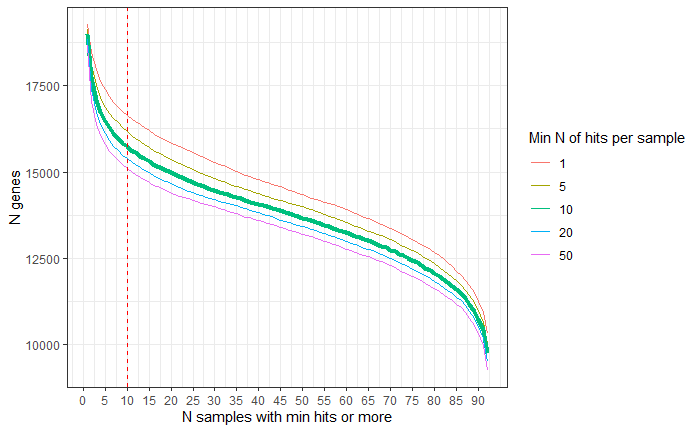
## DPP-IV and ACE2 signature analysis

*(Need to compare cytokines and chemokines between COVID No COVID; COVID with and without DM AND; then between COVID with DM and between NO COVID and DM)*

## Differential Gene Expression Analyses

In total, 57,142 protein coding and noncoding genes and were sequenced in 92 samples of the COVID and non-COVID patients. Note that 175 genes had more than one variant each. For these genes, the counts of the variants were added up within each sample. Out of these, 19,907 were protein coding genes.

Genes with small number of hits were filtered out. Based on an examination of the number of genes remaining after filtering using varying the minimum number of hits per sample and the minimum number of samples with at least that many hits (Figure 1), both thresholds were set to 10. Hence, 14,223 genes with at least 10 hits in at least 10 out of 92 samples were used in the analysis.



Next, we identify coding genes differentially expressed in COVID patients compared to controls (at least 2-fold change and adjusted p-value ≤ 0.05). 2 genes were downregulated and 16 upregulated in COVID patients compared to non-COVID. The list of the 18 genes and the estimates of the differences (on log2 scale and representing the number of 2-fold changes in COVID vs. non-COVID patients’ samples) is presented in Table 3. Figure 1 shows the total number of hits for each of the 18 genes in each sample.



Figure 1: differentially expressed genes in COVID vs non-COVID

One of the most striking differences found in this part of the analysis was the overexpression of Interferon Alpha Inducible Protein 27 (IFI27) coding gene, with a 2-fold difference of 2.37 (SEM = 0.49). The protein coded by IFI27 was previously shown to be associated with viral infections including Hepatitis C, respiratory syncytial virus (RSV) infection and Enterovirus 71 (EV71) hand foot and mouth disease. It has been proposed as a biomarker for an early prediction of COVID-19 outcomes ([Shojaei et al 2022](https://pubmed.ncbi.nlm.nih.gov/36685600/)). In this study, most of the patients that had gene sequencing data and died in hospital had their IFI27 expression levels elevated compared to those who were discharged alive (Figure 2 left). No obvious patterns were observed for critical care as an outcome (Figure 2 right).



Figure 2: IFI27 expression in COVID vs. non-COVID color-coded for in-hospital deaths (left) and critical care (right). The only patient who died in the hospital and had an RNA-seq results had IFI27 levels that exceeded all the non-COVID patients' readouts.

Reactome Pathway Analysis applied to the differentially expressed18 genes revealed that

NOTE for Lu and Marshall: please check if my conclusions above make clinical sense, perhaps find more related studies showing IFI27 to be important for COVID, and what would be the mechanism of action. Also, are any of the rest of these genes are clinically interesting?

Next, DMs association with gene expressions was tested. In total, 431 genes were differentially expressed in DM patients vs. the controls (non-DM), with 45 genes upregulated by DM and 386 genes downregulated (Table 4.)

NOTE for Lu and Marshall: please check these genes for biological relevance for DM.

Out of all genes found to be differentially expressed in COVID/non-COVID and DM/non-DM patients, there were 5 genes in common: GRASP, KRT8, MYZAP, PRKG1 and SMIM24. The number of hits in the samples, grouped by COVID and DM diagnoses, are presented in Figure 3.



Figure 3: number of hits in samples, grouped by COVID and DM diagnoses.

For Davit and Vahe (1/18/2024):

1. Compare in-hospital death COVID/non-COVID patients’ gene expressions. Maybe add CCU (in-hospital death or CCU).
2. Javier’s suggestion: use Javier’s DMR package to aggregate omics data but weight each dataset. This way, gene expressions will not overwhelm the analysis.
3. Look only at the genes that coded for proteins in the ELISA dataset. High gene expression should correspond to high protein production unless there is a delay, mRNA gets blocked, or there are multiple transcriptomes that do not correspond to the observed proteins.

# Discussion

Inflammatory mediators measured:

IFN- α, IFN- γ, IL-1β, IL-2, IL-4, IL-5, IL-6, IL-8 (CXCL8), IL10, IL-12p70, IL-17A (CTLA-8), IL-18, IP-10 (CXCL10), MCP-1 (CCL2), MIP-1α, MIP-1β, TNFα, and TNFβ

# References