Pre-existing moderate or severe anemia is a risk factor for post-viral clearance hospitalization in COVID-19 patients

Patrick Lenehan[1]*, Eshwan Ramudu[1]*, AJ Venkatakrishnan[1], Gabriela Berner[1], Reid McMurry[1], John C. O'Horo[2], Andrew D. Badley[2], William Morice II[2,3], John Halamka[2], Venky Soundararajan[1+]

[1] nference, One Main Street, Cambridge, MA 02142, USA
[2] Mayo Clinic, Rochester, MN 55902, USA
[3] Mayo Clinic Laboratories, Rochester, MN 55902, USA
* Joint first authors
+ Address correspondence to Venky Soundararajan (venky@nference.net)

Abstract

As the number of new and recovering cases continues to rise, it is increasingly important to understand the long-term impacts of COVID-19 beyond the time of active SARS-CoV-2 infection. However, the clinical and molecular biomarkers characterizing patients who are likely to experience longer lasting effects after clearing SARS-CoV-2 are not yet known. We have found that COVID-19 patients hospitalized after confirmed viral clearance tended to have lower hemoglobin and hematocrit measurements both during their SARS-CoV-2 positive intervals and during the one year prior to COVID-19 diagnosis compared to patients who have not been rehospitalized after viral clearance. Further, outright laboratory-based diagnosis of moderate or severe anemia was strongly enriched in the hospitalized cohort, suggesting that anemia pre-dating or concurrent with SARS-CoV-2 infection may predispose patients to long-term complications of COVID-19. Interventions which may mitigate anemia did not reduce the risk of post clearance hospitalization, although the efficacy of and patient compliance with these interventions could not be established. This study demonstrates the value of integrated large-scale EHR analyses and highlights the need for further research to determine whether the prevention or mitigation of anemia during SARS-CoV-2 infection may reduce the risk of suffering long-term complications of COVID-19.

Introduction

Since the first diagnosed case of COVID-19 in December 2019, over 50 million people have been infected with SARS-CoV-2 worldwide resulting in over 1.25 million deaths ([1]). While significant progress has been made in understanding the pathogenesis of COVID-19, including the rapid development and clinical testing of multiple vaccine candidates ([2], [3], [4], [5], [6], [7]) along with detailed characterizations of the SARS-CoV-2 entry receptor ACE2 ([8], [9], [10], [11], [12]), there are still few options available for effective treatment of patients with severe COVID-19. Further, as the pandemic has progressed, there have been reports of long-lasting effects of COVID-19 even in patients who did not experience a severe disease course during their active period of infection ([13], [14], [15]). However, the clinical, molecular, and/or demographic

biomarkers characterizing patients who are more likely to experience these lasting effects after clearing SARS-CoV-2 are not yet known.

The need to answer such questions during the rapidly evolving COVID-19 pandemic has emphasized the requirement for tools facilitating real-time analysis of patient data as it is obtained and stored in large electronic health records (EHR) systems. Specifically, clinical research efforts to understand the features defining COVID-19 patients, or subsets thereof, fundamentally require reliable systems that enable (1) conversion of unstructured information (e.g., patients notes written by healthcare professionals) into structured formats suitable for downstream analysis and (2) temporal alignment and integration of such unstructured data with the already structured information available in EHR databases (e.g., lab test results, disease diagnosis codes).

With these requirements in mind, we have previously reported the development of augmented curation methods that enable the rapid creation and comparison of defined cohorts of COVID-19 patients within a large EHR system ([16], [17], [18], [19], [20]). Here we expand on our prior textual sentiment-based analysis ([16]) to understand the clinical features of patients likely to experience lasting effects of COVID-19, and we find that lab tests corroborate our previous results suggesting that anemia and kidney malfunction during active SARS-CoV-2 infection may serve as biomarkers of patients who are more likely to be subsequently hospitalized after PCR-confirmed viral clearance.

Results

Definition of COVID-19 cohorts based on hospitalization after PCR-confirmed SARS-CoV-2 clearance

Using NLP-based extraction of phenotypes from a large EHR system, we previously reported that COVID-19 patients who are hospitalized after viral clearance (as assessed by RT-PCR) were more likely to experience anemia and acute kidney injury (AKI) in the year prior to their diagnosis and during their PCR-positive phase of COVID-19 compared to patients who were not re-hospitalized after clearance of SARS-CoV-2 ([16]). Here, we sought to assess whether diagnostic lab tests for AKI and anemia corroborate these phenotypic associations. As was previously described, we split the cohort of COVID-19 patients with confirmed viral clearance into two groups: (1) post-clearance hospitalized ("PCH"; n = 93) and (2) post-clearance non-hospitalized ("PCNH"; n = 173), where viral clearance was defined as two consecutive negative SARS-CoV-2 PCR tests following a positive test. A demographic summarization of these two cohorts is provided in **Table 1**.

To reduce imbalances between cohorts based on initial infection severity, our cohort inclusion criteria required patients in the PCNH cohort (n = 173) to have been admitted to the hospital during their index infection (i.e. while positive for SARS-CoV-2), while it did not require the same for patients in the PCH cohort (only 49 of 93 were hospitalized during index infection). To assess whether the severity of the index infection differed between these groups, we used ICU admission rates as a surrogate of clinical severity. When considering all patients, the rate of ICU admission was similar in the two cohorts (26/93 [28%] vs. 57/173 [33%]; p = 0.49; **Figure S1A**), but when considering only patients who were hospitalized during index infection, the PCH cohort had a significantly higher rate of ICU admission (26/49 [53%] vs. 57/173 [33%]; p = 0.01;

Figure S1B). Since the patients who were hospitalized during index infection will inherently provide more lab testing data, it will be important in our subsequent analyses to evaluate whether lab test features distinguishing these cohorts are independent of this difference in ICU admission rate.

Lab tests indicate that patients hospitalized after viral clearance are more likely to experience anemia and kidney dysfunction before and during their SARS-CoV-2 infection

After creating these two cohorts, we then compared a set of selected lab test results during two time windows: (1) the year prior to COVID-19 diagnosis and (2) the time during which each patient was SARS-CoV-2 positive according to their RT-PCR results. Given our previous EHR-based findings, we considered both anemia-related and kidney function lab tests including hemoglobin, hematocrit, estimated glomerular filtration rate (eGFR), serum creatinine, and serum blood urea nitrogen (BUN) levels (**Figure 1**).

For each patient, we first considered the median values of the given lab test over the designated interval. Histograms showing the number of measurements per patient in each time period for the selected tests are shown in **Figures S2-S3**. Consistent with our previous findings, we found that patients in the PCH cohort showed significantly lower median hemoglobin and hematocrit levels in both the pre-COVID and the SARS-CoV-2 positive phases (**Table 2**, **Figures 2A-D**). In the SARS-CoV-2 positive phase, median eGFR was lower and median BUN was higher in the PCH cohort, but there were no significant differences in renal function test results between the cohorts in the pre-COVID phase (**Table 2**, **Figure S4**).

We also tested whether extreme (i.e. minimum or maximum) values of a given lab test over the designated periods varied between hospitalized and non-hospitalized patients, as a measure of central tendency (e.g., median) may fail to capture a single occurrence of phenotypes such as anemia or AKI. Specifically, we compared the patient-level minimum values of hemoglobin, hematocrit, and eGFR, and maximum values of serum creatinine and BUN in each time period. Interestingly, we found that the PCH cohort tended to have lower minimum values of hemoglobin, hematocrit, and eGFR in both the pre-COVID and SARS-CoV-2 positive phases, while also showing higher maximum serum BUN and creatinine during the SARS-CoV-2 positive phase (**Tables 3-4**, **Figure 3 and Figure S5**).

To consider a simple explanation, we asked whether these reductions in hemoglobin and hematocrit during the SARS-CoV-2 positive phase could be attributed to more frequent blood draws or higher ICU admission rates in PCH patients compared to their PCNH counterparts, but this was not the case. Specifically, there was no significant difference in the number of blood draws experienced by these cohorts during this interval (**Figure S6A**). And while patients admitted to the ICU during SARS-CoV-2 index infection tended to have slightly lower hemoglobin values compared to non-ICU admitted patients (cohen's D = -0.283, p = 0.022; **Figure S7**), the difference was much stronger when comparing the PCH and PCNH cohorts (cohen's D = -0.70, $p = 7.51 \times 10^{-6}$; **Table 2**). This indicates that the observed reduction in hemoglobin among PCH patients during the SARS-CoV-2 positive interval cannot be explained simply by their higher ICU admission rate during index infection (**Figure S1B**). Taken together, these analyses corroborate our prior textual sentiment-based EHR findings, suggesting that patients who are hospitalized after SARS-CoV-2

clearance are more likely to experience anemia and kidney malfunction both prior to and during SARS-CoV-2 infection.

Both male and female rehospitalized COVID-19 patients have lower hemoglobin and hematocrit measurements compared to patients who are not rehospitalized

Given that males and females have different normal ranges of hemoglobin and hematocrit, it was important to consider whether our observation was related to an imbalance in sex distributions between the PCH and PCNH cohorts. As seen in **Table 1**, this was especially important as the PCH cohort was indeed enriched for female patients (57/93 [61.3%]) compared to the PCNH cohort (71/173 [41%]).

To address this concern, we repeated our analysis of anemia lab tests split by sex. We found that patient-level median values of both hemoglobin and hematocrit during the SARS-CoV-2 positive phase were still significantly lower in both the male and female PCH cohorts versus their PCNH counterparts (**Table 5**, **Figures 4A-D**). Further, the pre-COVID median measurements of hematocrit and hemoglobin were lower in the female PCH cohort, although this was not the case among males (**Table 5**, **Figures 4E-H**). Similarly, in our analysis of extreme values, both male and female PCH patients showed lower minimum measurements of hemoglobin and hematocrit during the SARS-CoV-2 positive phase (**Table 6**, **Figures 4I-L**), while only females showed lower measurements in the pre-COVID phase (**Table 6**, **Figures 4M-P**). We again confirmed that the reductions observed during the SARS-CoV-2 positive phase were not attributable to more frequent blood draws, as there was no significant difference in the number of blood draws per patient between these cohorts when split by sex (**Figures S2B-C**).

Rehospitalized patients are more likely to have experienced moderate or severe anemia prior to COVID-19 diagnosis and during active COVID-19 infection

The previous analyses have relied simply on the comparison of continuous lab test measurements between two cohorts without clinical interpretations of those measurements. However, given that anemia can be defined based on the lab tests we have considered, we next sought to determine whether the outright diagnosis of anemia occurs more frequently in the PCH cohort than the PCNH cohort. To do so, we identified every case of clinical anemia (defined as hemoglobin < 13.5 g/dL or hematocrit < 38.3% for males, and hemoglobin < 12.0 g/dL or hematocrit < 35.5% for females) for each patient in the pre-COVID and SARS-CoV-2 positive phases. Patients were then classified in a binary fashion for each time window based on whether the median of their pre-COVID phase measurements or the minimum of their SARS-CoV-2 positive phase measurements met the above criteria for laboratory-diagnosed anemia (defined as “pre-COVID anemia” and “SARS-CoV-2 positive anemia”, respectively) (**Figure 5A**). We found that anemia was observed more frequently in the PCH cohort during both the pre-COVID phase (27/42 [64%] vs. 23/65 [35%]; OR = 1.82; p = 0.005) and the SARS-CoV-2 positive phase (53/62 [85%] vs. 123/167 [74%]; OR = 1.16; p = 0.077), although the latter finding did not reach statistical significance (**Figures 5B-C**).

Anemia can be further characterized as mild, moderate, or severe based on hemoglobin measurements, where a hemoglobin level <10 g/dL is consistent with moderate or severe disease for both males and females. We thus performed a similar analysis to the one described above,

but classified patients in a binary fashion for each time window based on whether the median of their pre-COVID phase hemoglobin measurements or the minimum of their SARS-CoV-2 positive phase measurements was <10g/dL. Here we found that moderate/severe anemia was observed more frequently in the PCH cohort during both the pre-COVID phase (7/42 [17%] vs. 2/65 [3%]; OR = 5.42; p = 0.027) and the SARS-CoV-2 positive phase (33/62 [53%] vs. 36/167 [22%]; OR = 2.47; p = 8.2×10^{-6}) (**Figures 5D-E**).

Again it was important to assess whether this strong enrichment for moderate/severe anemia in PCH patients during the SARS-CoV-2 positive phase was related to or independent of index infection severity as estimated by ICU admission status. We found that among patients admitted to the ICU during index infection, the prevalence of moderate or severe anemia in the PCH cohort was 77% (20/26) compared to 30% (16/53) in the PCNH cohort (OR = 2.55; p = 1.11×10^{-4}; **Figure S8A**); among those not admitted to the ICU during index infection, the prevalence of moderate or severe anemia was 36% (13/36) in the PCH cohort versus 18% (20/114) in the PCNH cohort (OR = 2.06; p = 0.036 **Figure S8B**). Thus, it seems that the occurrence of moderate or severe anemia during the SARS-CoV-2 positive interval is strongly associated with post-clearance hospitalization status regardless of index infection severity.

Given these findings, we evaluated whether therapeutic interventions to treat moderate or severe anemia prior to or during COVID-19 infection provided any protective effect against post clearance hospitalization. We classified all patients with moderate or severe anemia on the basis of whether they received any anemia-targeted intervention during the respective time interval. Anemia-targeted interventions included vitamin or mineral supplementation (iron, vitamin B12, multivitamins), recombinant erythropoietin, and red blood cell transfusions. We found that patients who were administered one or more of these interventions were hospitalized after viral clearance at similar rates to patients who did not receive them in both the pre-COVID phase (5/7 [71%] vs. 2/2 [100%]; OR = 0.71; p = 1.0) and SARS-CoV-2 positive phase (26/52 [50%] vs. 7/17 [41%]; OR = 1.091; p = 0.59) (**Figures 5F-G**). It is important to note that both of these analyses (especially that of the pre-COVID phase) were limited due to the small total number of patients with moderate or severe anemia in these time periods (n = 9 for pre-COVID, 69 for SARS-CoV-2 positive).

Taken together, we conclude the following about COVID-19 patients who are hospitalized after viral clearance (PCH patients) compared to those who are not (PCNH patients): (1) they tend to have lower hemoglobin and hematocrit levels in the year prior to COVID-19 diagnosis and during active COVID-19 infection, (2) they are more likely to be diagnosed with moderate or severe anemia during both intervals; and (3) anemia-mitigating interventions in the hospital setting do not appear to reduce the risk of post clearance hospitalization.

Prior diagnosis of acute kidney injury is not associated with rehospitalization status in COVID-19 patients

Given that AKI is not defined by static lab measurements, but rather by changes in measurements (e.g. serum creatinine) over time, we next classified patients on the basis of whether they experienced laboratory-confirmed AKI during the pre-COVID and/or SARS-CoV-2 positive intervals. Specifically, we referred to the creatinine-related components of the KDIGO (Kidney Disease: Improving Global Outcomes) criteria for diagnosis and staging of AKI in adults:

stage 1 AKI is characterized by an increase in serum creatinine by ≥0.3 mg/dL within 48 hours or an increase in serum creatinine to ≥1.5x the baseline value (known or assumed to have occurred within the last 7 days); stage 2 AKI is characterized by an increase in serum creatinine to ≥2x the baseline value; and stage 3 AKI is characterized by an increase in serum creatinine to ≥3x the baseline value or to ≥4 mg/dL (**Figure S9A**).

We then used our longitudinal lab testing data to identify all instances meeting these criteria for each patient, and we compared the number of patients in the PCH and PCNH cohorts who did and did not experience AKI during the pre-COVID phase and during the SARS-CoV-2 positive phase. This analysis showed that similar fractions of patients in the two cohorts experienced any stage AKI during the pre-COVID phase (2/28 [7%] vs 6/53 [11%]; OR = 0.63; p = 0.71) (**Figure S9B**). In the SARS-CoV-2 positive phase, the rate of AKI was slightly higher in the PCH cohort, but this finding was not statistically significant (15/38 [39%] vs. 23/92 [25%]; OR = 1.58; p = 0.14) (**Figure S9C**).

Similar to our previous anemia analysis, we also asked whether the rate of moderate/severe AKI varied between these two cohorts. Here we found that stage 2+ AKI was slightly more common in the PCH cohort during both the pre-COVID (4/28 [14%] vs. 3/53 [6%]; OR = 2.52; p = 0.23) and SARS-CoV-2 positive (6/38 [16%] vs. 7/92 [8%]; OR = 2.08; p = 0.20) phases, but again these associations were not statistically significant (**Figure S9D-E**). Intriguingly when split by sex, we found that male PCH patients were more likely to experience stage 2+ AKI in both the pre-COVID (3/14 [21%] vs. 1/31 [3%]; OR = 6.64; p = 0.08) and SARS-CoV-2 positive (5/19 [26%] vs. 1/53 [2%]; OR = 13.95; p = 0.004) phases (**Figure S9F-G**), but this analysis is limited by the total cohort sizes (n = 45 for pre-COVID comparison; n = 72 for SARS-CoV-2 positive comparison).

Taken together, these results suggest that laboratory-based diagnosis of AKI in the year prior to COVID-19 diagnosis is not associated with post viral clearance hospitalization, while the diagnosis of stage 2+ AKI during the SARS-CoV-2 positive phase may be associated with subsequent hospitalization in male patients. However, the validity of this association must be further tested in larger patient cohorts.

Discussion

Almost one year after the first confirmed case, the COVID-19 pandemic continues to ravage communities across the globe. While efforts early in the pandemic rightly focused on the acute lung inflammation caused by SARS-CoV-2, the subsequent realization that COVID-19 may have more lasting effects has mandated a better understanding of factors that predispose patients to experience long-term COVID-19 related complications. We have previously sought to address this knowledge gap using state-of-the-art NLP models deployed on a massive EHR system, and here we have expanded this effort to include the longitudinal analysis of laboratory measurements both prior to COVID-19 diagnosis and during active SARS-CoV-2 infection.

It is first important to note that this study has several limitations. First, this analysis considers only patients within one EHR system; while this system does contain patient data from multiple sites of clinical care in distinct geographic locations (Minnesota, Arizona, Florida), there are still likely underlying biases in important factors such as patient demographics and tendencies around the ordering of laboratory tests by clinicians. Such biases could prevent the studied cohort

and their associated data points from serving as true representative samples of all COVID-19 patients. Second, as was previously mentioned, the analyzed cohort was relatively small, with even smaller patient counts available for particular lab tests of interest. Finally, we defined a SARS-CoV-2 positive window based on the presence of two negative RT-PCR tests, but the likely true date of clearance would precede the first negative test by an unknown amount of time.

Consistent with our previous conclusions, this lab test analysis suggests that both anemia and renal function in the pre-COVID and SARS-CoV-2 positive phases are associated with the risk of post viral clearance hospitalization. While the pathophysiologic basis for these associations are not yet clear, the findings do merit consideration in the context of clinical care of COVID-19 patients. Indeed, pre-existing conditions are already integrated in the clinical decision-making algorithms around COVID-19 as the Center for Diseases Control (CDC) has designated various chronic conditions as risk factors for severe COVID-19 infection (e.g. cancer, chronic kidney disease, chronic obstructive pulmonary disease, and cardiovascular diseases such as heart disease, obesity, and diabetes). However, there is much less known regarding factors or conditions that place people at risk for subsequent complications such as rehospitalization after viral clearance. Once identified, such factors and conditions should similarly be incorporated into the clinical decision-making process when treating COVID-19 patients.

Our finding that lower hemoglobin and hematocrit levels, and the outright diagnosis of moderate or severe anemia, prior to or during active SARS-CoV-2 infection is associated with post viral clearance hospitalization has not been previously reported. And while sickle cell disease is considered a risk factor for severe COVID-19, anemia itself is not considered to be such a risk factor. While we did not find evidence that administration of potential anemia-mitigating interventions was associated with lower risks of hospitalization after viral clearance, this does not rule out a role for anemia in long-term COVID-19 complications. It is important to note that we were not able to account for patient compliance in our analysis, and it is likely that patients who were moderately or severely anemic during their COVID-19 hospitalization were variably compliant with supplementation after their initial discharge. Further, even if mitigation of anemia does not impact subsequent hospitalization, these strong associations between anemia and COVID-19 are interesting in light of several previous lines of research.

First, several groups have reported an association between blood groups and susceptibility to and/or severity of COVID-19 infection ([21], [22], [23], [24]), suggesting that individuals with type O blood may be at lower risk for contracting COVID-19 or experiencing respiratory failure in the context of COVID-19. Whether this association reflects a direct or indirect interaction between SARS-CoV-2 and erythrocytes is not known, but it could certainly be relevant to pursue whether blood type is also associated with the occurrence or severity of anemia in the setting of COVID-19.

Second, fatigue has been commonly reported as both an acute symptom and a lasting effect of COVID-19 ([25], [26], [27]), but the mechanisms underlying this phenotype have not been established. It is worth noting that 179 of the 229 (78%) COVID-19 patients in this study had at least a mild anemia during their SARS-CoV-2 positive phase, and 103 of the 229 (45%) patients had a moderate or severe anemia (defined as hemoglobin < 11 g/dL) during this interval. It would be worthwhile to perform a longitudinal follow-up on these patients to determine whether they

continue to experience anemia in the months following SARS-CoV-2 clearance, and whether the presence of such a post-COVID anemia is associated with reports of fatigue.

Our findings regarding renal function tests and acute kidney injury may also be of clinical interest. Indeed, chronic kidney disease (CKD) has been recognized as a risk factor for severe COVID-19 infection. The fact that both median and extreme lab measurements suggest poorer renal function in the post clearance hospitalized cohort is consistent with this established risk factor, and suggests that the severity of one's CKD may have implications for their likelihood of hospitalization after viral clearance. Our finding that moderate/severe AKI during the SARS-CoV-2 positive phase was more frequently observed in the male PCH cohort compared to the male PCNH patients is interesting, but the small sample size available for analysis here would require further validation of this association.

Along with our previous analysis ([16]), this study illustrates the value of deploying sophisticated platforms across EHR systems that enable the integrated analysis of diverse data types including sentiment-laden text and laboratory test measurements. Taken together, these studies provide the first example of leveraging augmented curation methods to first identify phenotypes that distinguish defined clinical cohorts and to then cross check these phenotypic associations through a hypothesis-driven analysis of the most relevant lab tests. This framework can be effectively scaled for other clinical research efforts not only in COVID-19 but also in any disease areas of interest.

Methods

Study design

This was a nested case-control study evaluating 93 post SARS-CoV-2 clearance hospitalized patients and 173 post SARS-CoV-2 clearance non-hospitalized patients, selected from a cohort of 22,223 patients who presented to the Mayo Clinic Health System (including tertiary medical centers in Minnesota, Arizona, and Florida) and received at least one positive SARS-CoV-2 PCR test between the start of the COVID-19 pandemic and October 27, 2020 (see **Figure 1B**). This retrospective research was conducted with approval from the Mayo Clinic Institutional Review Board (IRB 20–003278).

Selection of study participants

The definitions of the cases and controls were described in detail previously ([16]). Briefly, post clearance hospitalized ("PCH") cases (N=93) were defined as patients who had two documented negative SARS-CoV-2 PCR tests following their last positive test result and were subsequently admitted to the hospital within 90 days of clearance. Post clearance non-hospitalized ("PCNH") controls (N=173) were defined as those who had two documented negative SARS-CoV-2 PCR tests following their last positive test result and were not hospitalized within 90 days of clearance. As shown in **Table 1**, there were no statistically significant differences between these groups in age, relative cleared date (defined as time to second negative SARS-CoV-2 PCR test after first positive test), race, and ethnicity. A majority of hospitalized post-clearance cases were female as compared to non-hospitalized controls (61.3% vs. 41.0%, p <.01).

Exposures and outcomes measures

Definition of the considered time intervals: pre-COVID phase and SARS-CoV-2 positive phase

Laboratory results were assessed during the year prior to COVID-19 diagnosis, hitherto referred to as the pre-COVID-phase, and during the period in which SARS-CoV-2 PCR was positive, hitherto referred to as the SARS-CoV-2 positive phase. A diagnosis of COVID-19 was conferred by a positive SARS-CoV-2 RT-PCR test and clearance was defined as two consecutive negative SARS-CoV-2 RT-PCR tests occurring after a positive test.

Selection and summarization of laboratory measurements

Our primary outcomes were anemia and AKI. Laboratory measurements related to anemia included hemoglobin and hematocrit, and laboratory measurements related to AKI included serum creatinine, serum blood urea nitrogen (BUN), and estimated glomerular filtration rate (eGFR). For a given lab test, we considered the median, maximum, and minimum measurements for each patient during the specified time windows (i.e., one year prior to COVID-19 diagnosis, and the PCR-defined SARS-CoV-2 positive phase). Given the directionality of these tests (i.e. anemia is defined by low hemoglobin and hematocrit, while kidney dysfunction is characterized by increases in serum creatinine and BUN but a decrease in eGFR), we were primarily interested in comparing the patient-level minimum values of hemoglobin, hematocrit, and eGFR, and maximum values of serum creatinine and BUN in each time period.

Classification of patients using clinical diagnostic criteria for anemia and AKI

We also classified patients in a binary fashion for each time window based on whether their lab tests were consistent with the clinical diagnosis of anemia or acute kidney injury. Classifications were defined according to the Mayo Clinic reference ranges for anemia ([28], [28,29]) and the KDIGO (Kidney Disease: Improving Global Outcomes) criteria for AKI ([28–30]) as follows:

- *Pre-COVID anemia (mild, moderate, or severe)*: for males, median hemoglobin < 13.5 g/dL or median hematocrit < 38.3% during the one year prior to PCR confirmation of COVID-19 diagnosis. For females, median hemoglobin < 12.0 g/dL or median hematocrit < 35.5% during the one year prior to PCR confirmation of COVID-19 diagnosis.
- *SARS-CoV-2 positive anemia (mild, moderate, or severe)*: for males, minimum hemoglobin < 13.5 g/dL or minimum hematocrit < 38.3% during the SARS-CoV-2 positive interval defined by PCR testing. For females, minimum hemoglobin < 12.0 g/dL or minimum hematocrit < 35.5% during the SARS-CoV-2 positive interval defined by PCR testing.
- *Pre-COVID anemia (moderate or severe)*: for both males and females, median hemoglobin < 10.0 g/dL during the one year prior to PCR confirmation of COVID-19 diagnosis.
- *SARS-CoV-2 positive anemia (moderate or severe)*: for both males and females, minimum hemoglobin < 10.0 g/dL during the SARS-CoV-2 positive interval defined by PCR testing.

- *Acute kidney injury (stage 1, 2, or 3)*: increase in serum creatinine by ≥0.3 mg/dL within 48 hours or an increase in serum creatinine to ≥1.5x the baseline value which is known or assumed to have occurred in the prior 7 days. Here, the baseline was defined as the minimum value among all serum creatinine tests for the given patient in the prior 7 days.
- *Acute kidney injury (stage 2 or 3)*: increase in serum creatinine to ≥2x the baseline value which is known or assumed to have occurred in the prior 7 days, or a serum creatinine value of ≥4 mg/dL. Again, the baseline was defined as the minimum value among all serum creatinine tests for the given patient in the prior 7 days.

Classification of patients based on administration of potential anemia mitigating interventions
To investigate the effects of anemia-targeted interventions in the year prior to or during COVID-19 infection, we identified patients from this cohort who received (1) blood transfusions per medical procedure documentation and/or (2) one or more of the following therapies per medication administration records: oral or intravenous iron, multivitamin or prenatal vitamins containing iron, oral or intramuscular vitamin B12 (cyanocobalamin), and intravenous or subcutaneous darbepoetin or erythropoietin.

Quantification of number of blood draws per patient
To test whether trends in anemia-related measurements could be explained by differences in the number of blood draws received in the pre-COVID or SARS-CoV-2 positive phases, we counted the number of blood draws in these time intervals for each patient. All tests with a documented source of “Blood”, “Plasma”, or “Serum” were collected for each patient; for a given patient on a given day, we then took the count of the most frequently obtained test as the number of blood draws for that patient on that day. For example, if the record for Patient *P* on Day *D* contained 5 serum sodium measurements, 3 hemoglobin measurements, and 1 plasma IL-6 measurement, then we inferred that Patient *P* received 5 blood draws on Day *D*.

Statistical Analysis

Laboratory values are reported as medians, minima, or maxima with interquartile range. Mann-Whitney U-tests were applied to continuous outcome measures, generating a *p*-value. Fisher exact tests were applied to categorical outcome measures, generating a *p*-value and an Odds Ratio. These tests were applied using the SciPy package[31] in Python (version 3.5). For each set of comparisons performed, p-values were corrected using a Benjamini-Hochberg correction for multiple hypothesis testing.

Figure Legends

Figure 1. Schematic summarizing cohort creation and lab test analyses. (A) Time intervals (“phases”) were defined based on SARS-CoV-2 PCR testing results. The pre-COVID phase refers to the one year prior to the date of the first positive SARS-CoV-2 PCR test. The SARS-CoV-2 positive interval refers to the time between the first positive SARS-CoV-2 PCR test and the second consecutive negative PCR test. The post-viral clearance phase refers to the time after this second consecutive negative PCR test. (B) From the 1355 patients who were diagnosed by PCR with

COVID-19 and subsequently confirmed to have cleared SARS-CoV-2 with two consecutive negative tests, we created two cohorts: (1) patients who were hospitalized during their index infection and not hospitalized after confirmed viral clearance (n = 173), and (2) patients who were hospitalized at some point after confirmed viral clearance (n = 93).

Figure 2. Median (baseline) values of (A) hemoglobin and (B) hematocrit during the Pre-COVID phase and (C) hemoglobin and (D) hematocrit during the SARS-CoV-2-positive phase. Both median hemoglobin and median hematocrit are significantly lower in the hospitalized cohort than in the non-hospitalized one. Red shading indicates normal ranges for hemoglobin and hematocrit; as these ranges are lower for females than males, the shaded range here spans from the lower limit of normal for females (12 g/dL hemoglobin, 35.5% hematocrit) to the upper limit of normal for males (17.5 g/dL hemoglobin, 48.6% hematocrit).

Figure 3. Comparison of minimum values for anemia-related tests (hemoglobin and hematocrit) in the pre-COVID (A-B) and SARS-CoV-2 positive (C-D) intervals. In both intervals, the hospitalized cohort tends to have lower measurements of both hemoglobin and hematocrit. Red shading indicates normal ranges for hemoglobin and hematocrit; as these ranges are lower for females than males, the shaded range here spans from the lower limit of normal for females (12 g/dL hemoglobin, 35.5% hematocrit) to the upper limit of normal for males (17.5 g/dL hemoglobin, 48.6% hematocrit).

Figure 4. Sex-split analysis of anemia-related lab tests in the pre-COVID and SARS-CoV-2 positive phases. (A-H) Median values of hemoglobin and hematocrit in the SARS-CoV-2 positive phase (A-D) and pre-COVID phase (E-H), split to show female patients (A-B, E-F) and male patients (C-D, G-H) separately. (I-P) Minimum values of hemoglobin and hematocrit in the SARS-CoV-2 positive phase (I-L) and pre-COVID phase (M-P), split to show female patients (I-J, M-N) and male patients (K-L, O-P) separately. For each comparison, statistics shown include Cohen's D, BH-corrected Mann Whitney U test p-value, and the difference of medians between the two cohorts. Red shading indicates normal ranges for hemoglobin and hematocrit depending on sex. For females, the normal ranges are 12.0-15.5 g/dL and 35.5-44.9%, respectively; for males, the ranges are 13.5-17.5 g/dL and 38.3-48.6%, respectively.

Figure 5. (A) Schematic illustrating how patients are classified as having no anemia, mild anemia, or severe/moderate anemia in the pre-COVID and SARS-CoV-2 positive intervals. (B-E) Comparison of anemia frequency in the hospitalized and non-hospitalized cohorts during the pre-COVID and SARS-CoV-2 positive phases. Contingency tables show the counts of patients with and without anemia in the hospitalized and non-hospitalized cohorts. Below each contingency table, the associated odds ratio and Fisher Exact test p-value is shown. (B-C) Mild, moderate, or severe anemia is defined as hemoglobin < 13.5 g/dL or hematocrit < 38.3% for males, and hemoglobin < 12.0 g/dL or hematocrit < 35.5% for females. (D-E) Moderate or severe anemia is defined as hemoglobin < 10 g/dL for both males and females. (F-G) Among patients with moderate or severe anemia (hemoglobin < 10 g/dL), comparison of the rates of administration of potential anemia-mitigating interventions between the hospitalized and non-hospitalized cohorts in the pre-COVID (F) and SARS-CoV-2 positive (G) intervals.

Acknowledgements

We thank Murali Aravamudan for his thoughtful review and feedback on this manuscript. We also thank Andrew Danielsen, Jason Ross, Jeff Anderson, and Sankar Ardhanari for their support that enabled the rapid completion of this study.

Competing Interests

PL, ER, AV, GB, RM, and VS are employees of nference and have financial interests in the company. ADB is a consultant for Abbvie, is on scientific advisory boards for nference and Zentalis, is founder and President of Splissen therapeutics. One or more of the investigators associated with this project from the Mayo Clinic and the Mayo Clinic have a Financial Conflict of Interest in technology used in the research. The investigator(s) and Mayo Clinic may stand to gain financially from the successful outcome of the research. This research has been reviewed by the Mayo Clinic Conflict of Interest Review Board and is being conducted in compliance with Mayo Clinic Conflict of Interest policies.

Author Contributions

Patrick Lenehan: Formal analysis, Supervision, Investigation, Methodology, Writing - original draft, Writing - review and editing. Contributed equally with Eshwan Ramudu.

Eshwan Ramudu: Data curation, Formal analysis, Investigation, Methodology, Writing - original draft, Writing - review and editing. Contributed equally with Patrick Lenehan.

AJ Venkatakrishnan: Supervision, Investigation, Methodology, Writing - review and editing.

Gabriela Berner: Data curation, Investigation, Writing - review and editing.

Reid McMurry: Methodology, Investigation, Writing - review and editing.

John C. O'Horo: Resources, Methodology, Project administration, Writing - review and editing.

Andrew D. Badley: Conceptualization, Resources, Formal analysis, Supervision, Validation, Funding acquisition, Investigation, Project administration, Writing - review and editing.

William Morice II: Resources, Supervision, Methodology, Project administration, Writing - review and editing.

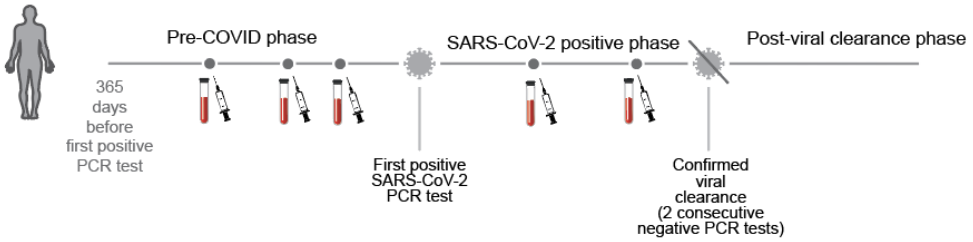
John Halamka: Resources, Supervision, Validation, Investigation, Methodology, Project administration, Writing - review and editing.

Venky Soundararajan: Conceptualization, Resources, Formal analysis, Supervision, Funding acquisition, Validation, Investigation, Methodology, Writing - original draft, Project administration, Writing - review and editing.

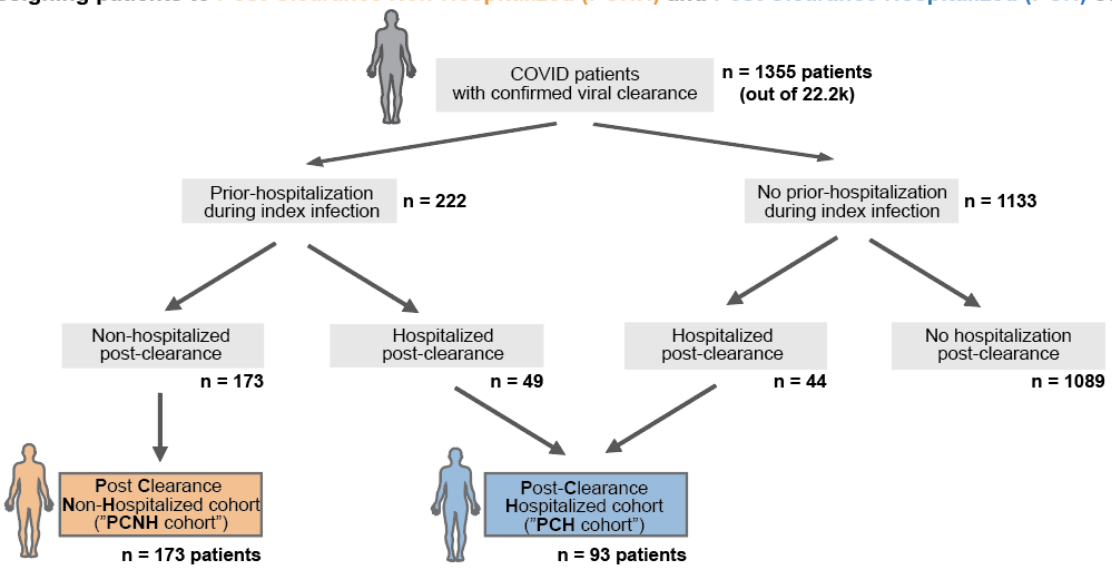
Data Availability

After publication, the de-deidentified data will be made available to others upon reasonable requests to the corresponding author and upon approval from the Mayo Clinic. A proposal with detailed description of study objectives and statistical analysis plan will be needed for evaluation of the reasonability of requests.

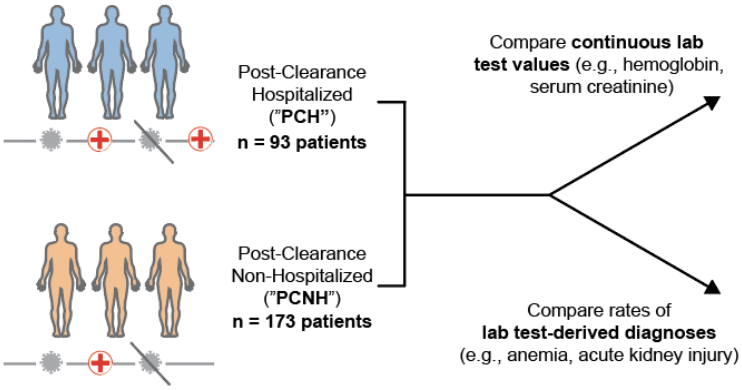
A Time-line capturing the clinical trajectory of a COVID patient based on PCR tests and lab tests

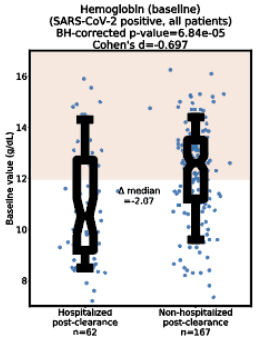


B Assigning patients to Post-Clearance Non-Hospitalized (PCNH) and Post Clearance Hospitalized (PCH) cohorts



C Comparing lab test results in defined cohorts during the pre-COVID and SARS-CoV-2 positive intervals





Moderate or Severe Anemia During SARS-CoV-2⁺ Phase

Hospitalized After Viral Clearance	Yes	No	Total
Yes	33	29	62
No	36	131	167
Total	69	160	229

Odds Ratio = 2.469
Fisher's Exact Test p-value = 8.18×10^{-6}

Figure 1

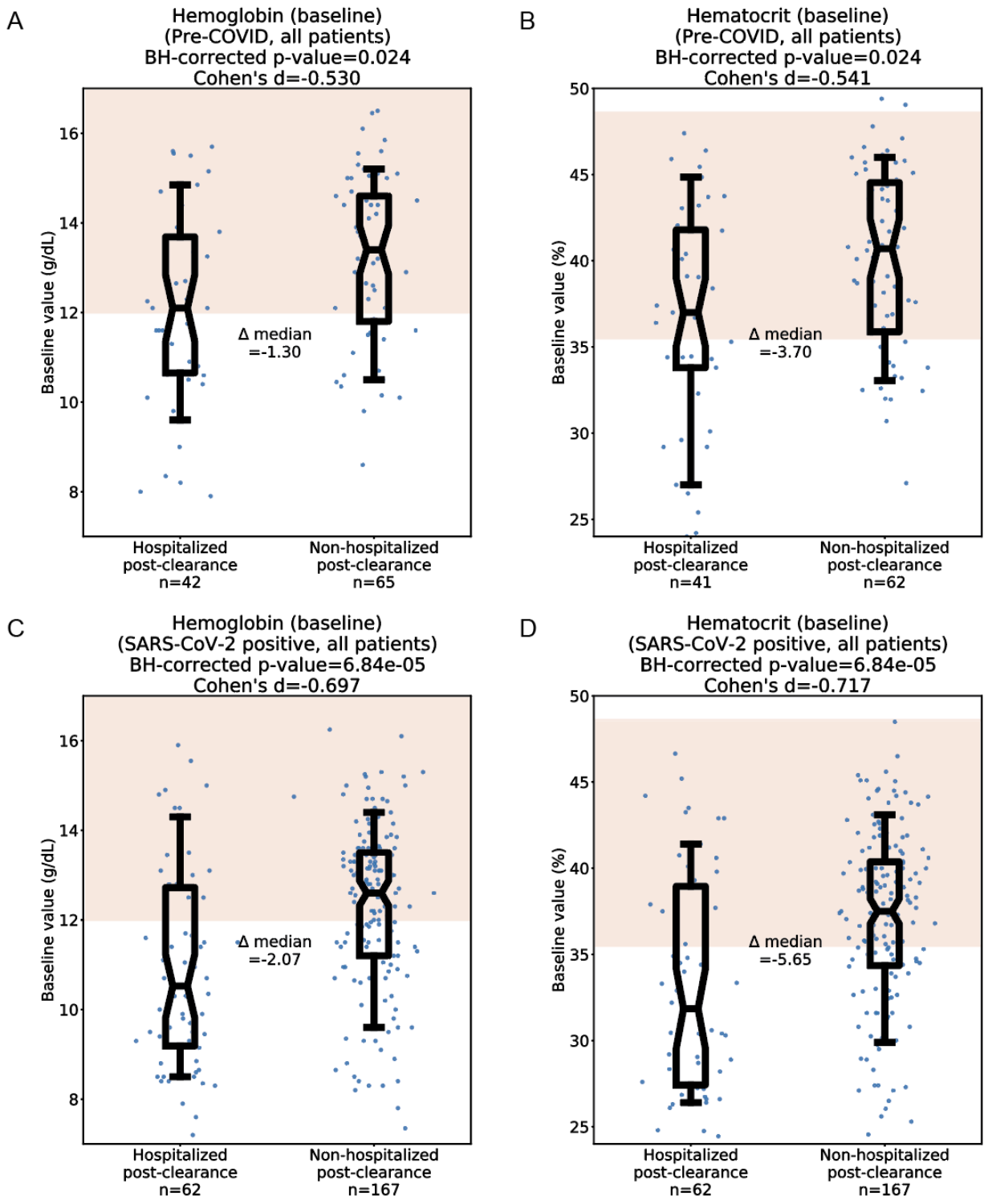

A
Hemoglobin (baseline)
(Pre-COVID, all patients)
BH-corrected p-value=0.024
Cohen's d=-0.530
Baseline value (g/dL)
Δ median
=-1.30
Hospitalized
post-clearance
n=42
Non-hospitalized
post-clearance
n=65
B
Hematocrit (baseline)
(Pre-COVID, all patients)
BH-corrected p-value=0.024
Cohen's d=-0.541
Baseline value (%)
Δ median
=-3.70
Hospitalized
post-clearance
n=41
Non-hospitalized
post-clearance
n=62
C
Hemoglobin (baseline)
(SARS-CoV-2 positive, all patients)
BH-corrected p-value=6.84e-05
Cohen's d=-0.697
Baseline value (g/dL)
Δ median
=-2.07
Hospitalized
post-clearance
n=62
Non-hospitalized
post-clearance
n=167
D
Hematocrit (baseline)
(SARS-CoV-2 positive, all patients)
BH-corrected p-value=6.84e-05
Cohen's d=-0.717
Baseline value (%)
Δ median
=-5.65
Hospitalized
post-clearance
n=62
Non-hospitalized
post-clearance
n=167

Figure 2

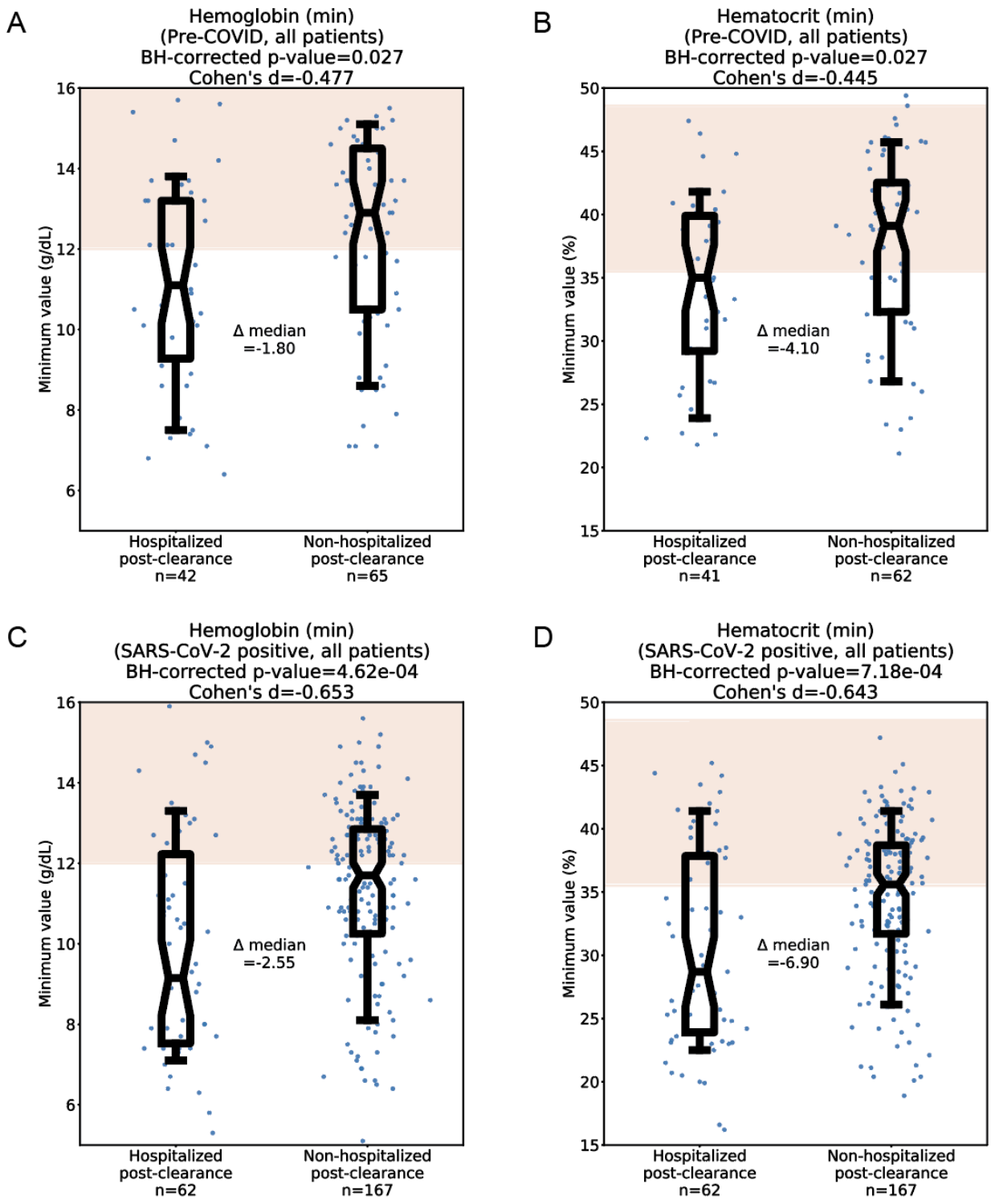

A
Hemoglobin (min)
(Pre-COVID, all patients)
BH-corrected p-value=0.027
Cohen's d=-0.477
Minimum value (g/dL)
Δ median
=-1.80
Hospitalized
post-clearance
n=42
Non-hospitalized
post-clearance
n=65
B
Hematocrit (min)
(Pre-COVID, all patients)
BH-corrected p-value=0.027
Cohen's d=-0.445
Minimum value (%)
Δ median
=-4.10
Hospitalized
post-clearance
n=41
Non-hospitalized
post-clearance
n=62
C
Hemoglobin (min)
(SARS-CoV-2 positive, all patients)
BH-corrected p-value=4.62e-04
Cohen's d=-0.653
Minimum value (g/dL)
Δ median
=-2.55
Hospitalized
post-clearance
n=62
Non-hospitalized
post-clearance
n=167
D
Hematocrit (min)
(SARS-CoV-2 positive, all patients)
BH-corrected p-value=7.18e-04
Cohen's d=-0.643
Minimum value (%)
Δ median
=-6.90
Hospitalized
post-clearance
n=62
Non-hospitalized
post-clearance
n=167

Figure 3

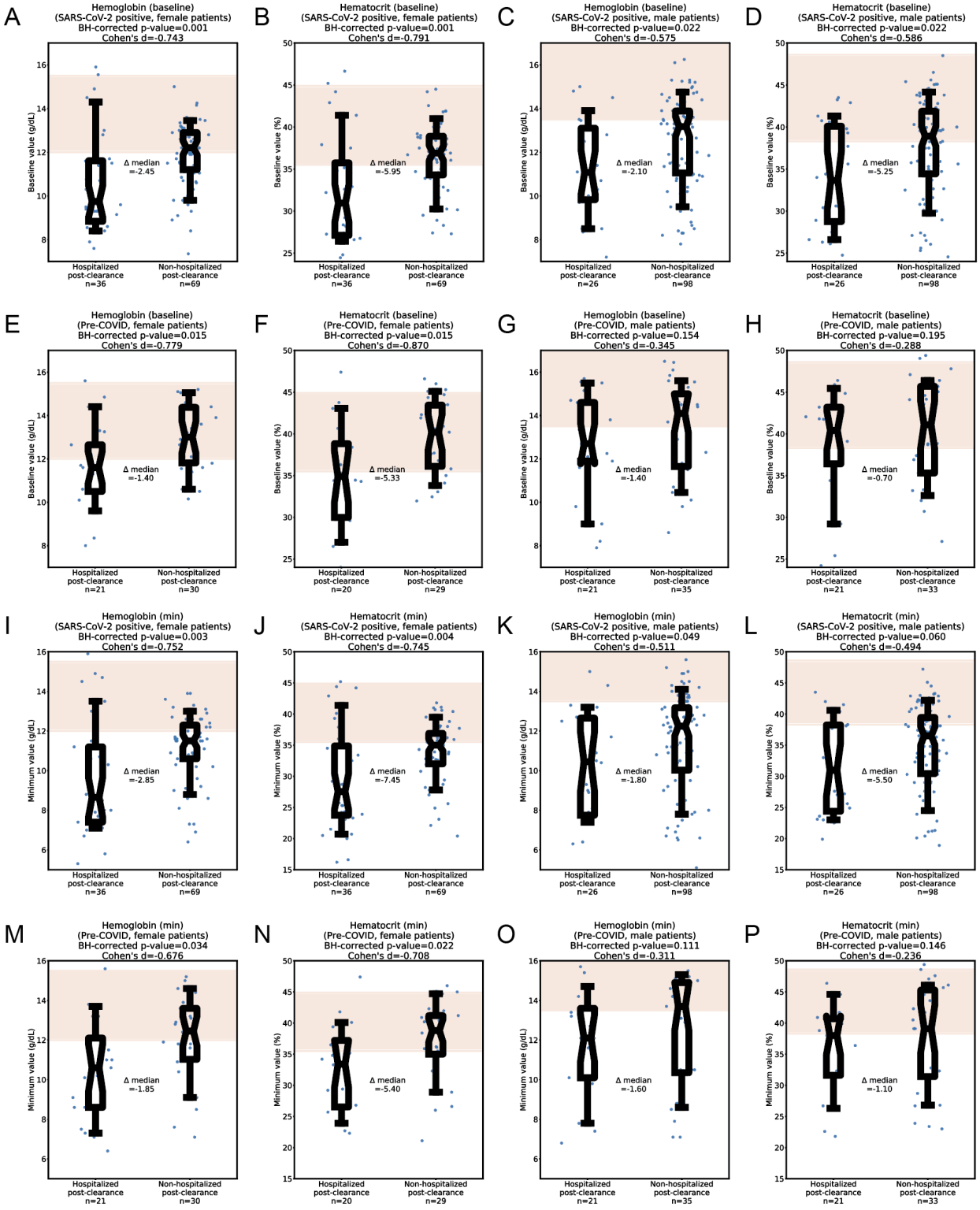

A
Hemoglobin (baseline)
(SARS-CoV-2 positive, female patients)
BH-corrected p-value=0.001
Cohen's d=-0.743
Baseline value (g/dL)
Δ median =-2.45
Hospitalized post-clearance n=36
Non-hospitalized post-clearance n=69
B
Hematocrit (baseline)
(SARS-CoV-2 positive, female patients)
BH-corrected p-value=0.001
Cohen's d=-0.791
Baseline value (%)
Δ median =-5.95
Hospitalized post-clearance n=36
Non-hospitalized post-clearance n=69
C
Hemoglobin (baseline)
(SARS-CoV-2 positive, male patients)
BH-corrected p-value=0.022
Cohen's d=-0.575
Baseline value (g/dL)
Δ median =-2.10
Hospitalized post-clearance n=26
Non-hospitalized post-clearance n=98
D
Hematocrit (baseline)
(SARS-CoV-2 positive, male patients)
BH-corrected p-value=0.022
Cohen's d=-0.586
Baseline value (%)
Δ median =-5.25
Hospitalized post-clearance n=26
Non-hospitalized post-clearance n=98
E
Hemoglobin (baseline)
(Pre-COVID, female patients)
BH-corrected p-value=0.015
Cohen's d=-0.779
Baseline value (g/dL)
Δ median =-1.40
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=30
F
Hematocrit (baseline)
(Pre-COVID, female patients)
BH-corrected p-value=0.015
Cohen's d=-0.870
Baseline value (%)
Δ median =-5.33
Hospitalized post-clearance n=20
Non-hospitalized post-clearance n=29
G
Hemoglobin (baseline)
(Pre-COVID, male patients)
BH-corrected p-value=0.154
Cohen's d=-0.345
Baseline value (g/dL)
Δ median =-1.40
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=35
H
Hematocrit (baseline)
(Pre-COVID, male patients)
BH-corrected p-value=0.195
Cohen's d=-0.288
Baseline value (%)
Δ median =-0.70
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=33
I
Hemoglobin (min)
(SARS-CoV-2 positive, female patients)
BH-corrected p-value=0.003
Cohen's d=-0.752
Minimum value (g/dL)
Δ median =-2.85
Hospitalized post-clearance n=36
Non-hospitalized post-clearance n=69
J
Hematocrit (min)
(SARS-CoV-2 positive, female patients)
BH-corrected p-value=0.004
Cohen's d=-0.745
Minimum value (%)
Δ median =-7.45
Hospitalized post-clearance n=36
Non-hospitalized post-clearance n=69
K
Hemoglobin (min)
(SARS-CoV-2 positive, male patients)
BH-corrected p-value=0.049
Cohen's d=-0.511
Minimum value (g/dL)
Δ median =-1.80
Hospitalized post-clearance n=26
Non-hospitalized post-clearance n=98
L
Hematocrit (min)
(SARS-CoV-2 positive, male patients)
BH-corrected p-value=0.060
Cohen's d=-0.494
Minimum value (%)
Δ median =-5.50
Hospitalized post-clearance n=26
Non-hospitalized post-clearance n=98
M
Hemoglobin (min)
(Pre-COVID, female patients)
BH-corrected p-value=0.034
Cohen's d=-0.676
Minimum value (g/dL)
Δ median =-1.85
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=30
N
Hematocrit (min)
(Pre-COVID, female patients)
BH-corrected p-value=0.022
Cohen's d=-0.708
Minimum value (%)
Δ median =-5.40
Hospitalized post-clearance n=20
Non-hospitalized post-clearance n=29
O
Hemoglobin (min)
(Pre-COVID, male patients)
BH-corrected p-value=0.111
Cohen's d=-0.311
Minimum value (g/dL)
Δ median =-1.60
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=35
P
Hematocrit (min)
(Pre-COVID, male patients)
BH-corrected p-value=0.146
Cohen's d=-0.236
Minimum value (%)
Δ median =-1.10
Hospitalized post-clearance n=21
Non-hospitalized post-clearance n=33

Figure 4

A

All patients with Defined Measurments in Interval of Interest
(Pre-COVID phase* or SARS-CoV-2 Positive Phase^)

*Median in Pre-COVID phase
^Minimum in SARS-CoV-2⁺ phase

Thresholds for
MALE and FEMALE patients

Hemoglobin (Hgb) in g/DL
Hematocrit (Hct) in %

Hgb > 13.5 AND Hct > 38.3	Hgb < 13.5 AND > 10 OR Hct < 38.3	Hgb < 10
Hgb > 12.0 AND Hct > 35.5	Hgb < 12 AND > 10 OR Hct < 35.5	Hgb < 10
NOT ANEMIC	MILD ANEMIA	MODERATE OR SEVERE ANEMIA

B

Anemia in Pre-COVID Phase (Mild, Moderate, or Severe)

Hospitalized	Yes	No	Total
Yes	27	15	42
No	23	42	65
Total	50	57	107

Odds Ratio = 1.817
Fisher's Exact Test p-value = 0.005

C

Anemia in SARS-CoV-2⁺ Phase (Mild, Moderate, or Severe)

Hospitalized	Yes	No	Total
Yes	53	9	62
No	123	44	167
Total	176	53	229

Odds Ratio = 1.161
Fisher's Exact Test p-value = 0.077

D

Anemia in Pre-COVID Phase (Moderate or Severe)

Hospitalized	Yes	No	Total
Yes	7	35	42
No	2	63	65
Total	9	98	107

Odds Ratio = 5.417
Fisher's Exact Test p-value = 0.027

E

Anemia in SARS-CoV-2⁺ Phase (Moderate or Severe)

Hospitalized	Yes	No	Total
Yes	33	29	62
No	36	131	167
Total	69	160	229

Odds Ratio = 2.469
Fisher's Exact Test p-value = 8.18×10^{-6}

F

Received Anemia Intervention (Vitamin/Mineral, EPO, Transfusion)

Hospitalized	Yes	No	Total
Yes	2	0	2
No	5	2	7
Total	7	2	9

Odds Ratio = 0.714
Fisher's Exact Test p-value = 1.00

G

Received Anemia Intervention (Vitamin/Mineral, EPO, Transfusion)

Hospitalized	Yes	No	Total
Yes	7	10	17
No	26	26	52
Total	33	36	69

Odds Ratio = 1.091
Fisher's Exact Test p-value = 0.585

Figure 5

Demographic or Clinical Characteristic	Hospitalized Post-Clearance	Non-Hospitalized Post-Clearance	BH-adjusted p-value
Total number of patients, *N*	93	173	
Age *years (standard deviation)*	55.1 (19.9)	58.5 (16.7)	0.22
Relative Cleared Date, *days (standard deviation)*	21.3 (13.5)	21.7 (12.8)	0.35
Hospitalized during index infection, *n (%)*	49 (52.7%)	173 (100%)	0.01
Sex			
- Female, *n (%)*	57 (61.3%)	71 (41.0%)	0.01
- Male, *n (%)*	36 (38.7%)	102 (59.0%)	0.01
Race			
- White, *n (%)*	61 (65.6%)	104 (60.1%)	0.51
- Asian, *n (%)*	4 (5.4%)	18 (10.4%)	0.22
- Black, *n (%)*	11 (12.9%)	22 (12.7%)	1.00
- Other, *n (%)*	15 (16.1%)	29 (16.8%)	1.00
Ethnicity			
- Hispanic, *n (%)*	14 (15.1%)	39 (22.5%)	0.26
- Non-Hispanic, *n (%)*	77 (82.8%)	127 (73.4%)	0.22
- Other, *n (%)*	1 (2.2%)	7 (4.0%)	0.36

Table 1. Demographics and clinical characteristics of study cohorts, including patients who were and who were not hospitalized after PCR-confirmed clearance of SARS-CoV-2. The “Hospitalized post-clearance” cohort is composed of patients who were admitted or readmitted to the hospital following the estimated clearance date of SARS-CoV-2 infection. The “non-hospitalized post-clearance” cohort is composed of patients who were not admitted to the hospital following the estimated clearance date. Each demographic variable or clinical characteristic was tested for difference in proportion with a Fisher Exact test or a difference in magnitude (for continuous variables) using a Mann-Whitney U test, and p-values were corrected for multiple testing using the Benjamini-Hochberg (BH) correction.

Lab Test	Units	Phase	N (PCH cohort)	Mean of Medians (PCH cohort)	N (PCNH cohort)	Mean of Medians (PCNH cohort)	Cohen's D	P-Value	BH-adj P-value
Hemoglobin	g/dL	SARS-CoV-2 positive	62	10.914	167	12.272	-0.697	7.511E-06	6.838E-05
Hematocrit	%	SARS-CoV-2 positive	62	33.033	167	37.007	-0.717	1.368E-05	6.838E-05
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	47	61.436	107	71.262	-0.507	1.875E-03	6.250E-03
Hemoglobin	g/dL	Pre-COVID	42	12.113	65	13.182	-0.530	1.010E-02	2.359E-02
Hematocrit	%	Pre-COVID	41	36.899	62	40.030	-0.541	1.179E-02	2.359E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	35	25.266	92	20.819	0.300	2.351E-02	3.919E-02
Creatinine	mg/dL	SARS-CoV-2 positive	38	1.214	92	1.142	0.057	1.017E-01	1.452E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	25	62.760	50	68.660	-0.355	1.389E-01	1.737E-01
Blood Urea Nitrogen	mg/dL	Pre-COVID	23	24.835	47	18.328	0.430	2.027E-01	2.252E-01
Creatinine	mg/dL	Pre-COVID	25	2.001	53	1.223	0.371	2.466E-01	2.466E-01

Table 2. Analysis of median values for all selected lab tests in pre-COVID and SARS-CoV-2 positive phases, including both male and female patients.

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Minimums (PCH Cohort)	N (PCNH Cohort)	Mean of Minimums (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Hemoglobin	g/dL	SARS-CoV-2 positive	62	9.866	167	11.372	-0.653	4.616E-05	4.616E-04
Hematocrit	%	SARS-CoV-2 positive	62	30.342	167	34.514	-0.643	1.436E-04	7.180E-04
Hemoglobin	g/dL	Pre-COVID	42	11.157	65	12.375	-0.477	9.673E-03	2.701E-02
Hematocrit	%	Pre-COVID	41	34.373	62	37.587	-0.445	1.128E-02	2.701E-02
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	47	50.000	107	59.748	-0.418	1.351E-02	2.701E-02
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	25	52.400	50	61.340	-0.471	2.487E-02	4.145E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	35	16.526	92	14.543	0.190	1.759E-01	2.513E-01
Blood Urea Nitrogen	mg/dL	Pre-COVID	23	19.561	47	15.447	0.335	2.616E-01	3.079E-01
Creatinine	mg/dL	SARS-CoV-2 positive	38	0.997	92	0.929	0.074	2.772E-01	3.079E-01
Creatinine	mg/dL	Pre-COVID	25	1.666	53	1.098	0.331	3.111E-01	3.111E-01

Table 3. Analysis of minimum values for all selected lab tests in pre-COVID and SARS-CoV-2 positive phases, including both male and female patients.

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Maximums (PCH Cohort)	N (PCNH Cohort)	Mean of Maximums (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Hemoglobin	g/dL	Pre-COVID	42	13.117	65	13.882	-0.439	3.909E-02	6.441E-02
Hemoglobin	g/dL	SARS-CoV-2 positive	62	12.431	167	13.594	-0.590	9.813E-05	4.907E-04
Hematocrit	%	Pre-COVID	41	39.978	62	42.202	-0.451	3.547E-02	6.441E-02
Hematocrit	%	SARS-CoV-2 positive	62	37.211	167	40.523	-0.616	6.836E-05	4.907E-04
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	25	71.120	50	75.000	-0.243	3.742E-01	3.742E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	47	69.468	107	80.383	-0.608	8.879E-04	2.960E-03
Blood Urea Nitrogen	mg/dL	Pre-COVID	23	33.887	47	22.894	0.450	7.683E-02	9.604E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	35	42.343	92	29.533	0.502	3.963E-03	9.908E-03
Creatinine	mg/dL	Pre-COVID	25	2.328	53	1.378	0.394	1.662E-01	1.847E-01
Creatinine	mg/dL	SARS-CoV-2 positive	38	1.759	92	1.500	0.137	4.508E-02	6.441E-02

Table 4. Analysis of maximum values for all selected lab tests in pre-COVID and SARS-CoV-2 positive phases, including both male and female patients.

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Medians (PCH Cohort)	N (PCNH Cohort)	Mean of Medians (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Hemoglobin	g/dL	SARS-CoV-2 positive	36	10.611	69	11.920	-0.743	1.049E-04	1.049E-03
Hematocrit	%	SARS-CoV-2 positive	36	32.312	69	36.374	-0.791	2.098E-04	1.049E-03
Hemoglobin	g/dL	Pre-COVID	21	11.593	30	12.925	-0.779	6.069E-03	1.517E-02
Hematocrit	%	Pre-COVID	20	35.130	29	39.645	-0.870	5.659E-03	1.517E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	17	20.129	39	19.033	0.069	1.012E-01	2.024E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	24	69.562	41	73.537	-0.226	1.239E-01	2.064E-01
Blood Urea Nitrogen	mg/dL	Pre-COVID	9	18.500	20	15.118	0.252	2.464E-01	3.226E-01
Creatinine	mg/dL	Pre-COVID	11	1.735	22	1.266	0.170	2.580E-01	3.226E-01
Creatinine	mg/dL	SARS-CoV-2 positive	19	0.784	39	1.403	-0.368	3.037E-01	3.375E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	11	73.227	18	73.750	-0.052	4.731E-01	4.731E-01

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Medians (PCH Cohort)	N (PCNH Cohort)	Mean of Medians (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	23	52.957	66	69.848	-0.854	6.193E-04	6.193E-03
Creatinine	mg/dL	SARS-CoV-2 positive	19	1.643	53	0.949	0.989	1.613E-03	8.063E-03
Hemoglobin	g/dL	SARS-CoV-2 positive	26	11.333	98	12.519	-0.575	6.794E-03	2.172E-02
Hematocrit	%	SARS-CoV-2 positive	26	34.031	98	37.453	-0.586	8.689E-03	2.172E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	18	30.117	53	22.133	0.585	1.296E-02	2.591E-02
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	14	54.536	32	65.797	-0.621	5.340E-02	8.901E-02
Blood Urea Nitrogen	mg/dL	Pre-COVID	14	28.907	27	20.706	0.519	8.041E-02	1.149E-01
Hemoglobin	g/dL	Pre-COVID	21	12.633	35	13.403	-0.345	1.231E-01	1.539E-01
Hematocrit	%	Pre-COVID	21	38.583	33	40.368	-0.288	1.758E-01	1.953E-01
Creatinine	mg/dL	Pre-COVID	14	2.210	31	1.192	0.676	2.864E-01	2.864E-01

Table 5. Sex-split analysis of median values for all selected lab tests in pre-COVID and SARS-CoV-2 positive phases. Data is shown separately for females (top) and males (bottom).

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Minimums (PCH Cohort)	N (PCNH Cohort)	Mean of Minimums (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Hemoglobin	g/dL	SARS-CoV-2 positive	36	9.583	69	11.177	-0.752	2.625E-04	2.625E-03
Hematocrit	%	SARS-CoV-2 positive	36	29.528	69	34.068	-0.745	7.990E-04	3.995E-03
Hematocrit	%	Pre-COVID	20	32.620	29	37.259	-0.708	6.535E-03	2.178E-02
Hemoglobin	g/dL	Pre-COVID	21	10.567	30	12.103	-0.676	1.351E-02	3.377E-02
Creatinine	mg/dL	SARS-CoV-2 positive	19	0.653	39	1.094	-0.375	6.263E-02	1.253E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	11	57.727	18	65.667	-0.507	8.497E-02	1.416E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	24	57.000	41	61.512	-0.198	2.132E-01	2.800E-01
Blood Urea Nitrogen	mg/dL	Pre-COVID	9	15.722	20	13.980	0.125	2.240E-01	2.800E-01
Creatinine	mg/dL	Pre-COVID	11	1.653	22	1.119	0.215	3.725E-01	4.139E-01
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	17	13.018	39	13.949	-0.073	4.327E-01	4.327E-01

Lab Test	Units	Phase	N (PCH Cohort)	Mean of Minimums (PCH Cohort)	N (PCNH Cohort)	Mean of Minimums (PCNH Cohort)	Cohen's D	P-Value	BH-Adj P-Value
Estimated Glomerular Filtration Rate	mL/min/BSA	SARS-CoV-2 positive	23	42.696	66	58.652	-0.684	4.215E-03	2.240E-02
Creatinine	mg/dL	SARS-CoV-2 positive	19	1.341	53	0.809	0.950	4.479E-03	2.240E-02
Hemoglobin	g/dL	SARS-CoV-2 positive	26	10.258	98	11.509	-0.511	1.476E-02	4.919E-02
Hematocrit	%	SARS-CoV-2 positive	26	31.469	98	34.829	-0.494	2.387E-02	5.968E-02
Blood Urea Nitrogen	mg/dL	SARS-CoV-2 positive	18	19.839	53	14.981	0.613	4.986E-02	9.973E-02
Hemoglobin	g/dL	Pre-COVID	21	11.748	35	12.609	-0.311	8.644E-02	1.111E-01
Estimated Glomerular Filtration Rate	mL/min/BSA	Pre-COVID	14	48.214	32	58.906	-0.520	8.110E-02	1.111E-01
Blood Urea Nitrogen	mg/dL	Pre-COVID	14	22.029	27	16.533	0.502	8.889E-02	1.111E-01
Hematocrit	%	Pre-COVID	21	36.043	33	37.876	-0.236	1.318E-01	1.464E-01
Creatinine	mg/dL	Pre-COVID	14	1.676	31	1.084	0.678	2.347E-01	2.347E-01

Table 6. Sex-split analysis of minimum values for all selected lab tests in pre-COVID and SARS-CoV-2 positive phases. Data is shown separately for females (top) and males (bottom).

References

1. COVID-19 Map - Johns Hopkins Coronavirus Resource Center. [cited 8 Dec 2020]. Available: https://coronavirus.jhu.edu/map.html
2. Website. [cited 8 Dec 2020]. Available: https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(20)31604-4/fulltext
3. Jackson LA, Anderson EJ, Rouphael NG, Roberts PC, Makhene M, Coler RN, et al. An mRNA Vaccine against SARS-CoV-2 - Preliminary Report. N Engl J Med. 2020;383. doi:10.1056/NEJMoa2022483
4. Corbett KS, Flynn B, Foulds KE, Francica JR, Boyoglu-Barnum S, Werner AP, et al. Evaluation of the mRNA-1273 Vaccine against SARS-CoV-2 in Nonhuman Primates. N Engl J Med. 2020;383: 1544–1555.
5. Mercado NB, Zahn R, Wegmann F, Loos C, Chandrashekar A, Yu J, et al. Single-shot Ad26 vaccine protects against SARS-CoV-2 in rhesus macaques. Nature. 2020;586: 583–588.
6. Sadoff J, Le Gars M, Shukarev G, Heerwegh D, Truyers C, de Groot AM, et al. Safety and immunogenicity of the Ad26.COV2.S COVID-19 vaccine candidate: interim results of a phase 1/2a, double-blind, randomized, placebo-controlled trial. medRxiv. 2020; 2020.09.23.20199604.
7. Mulligan MJ, Lyke KE, Kitchin N, Absalon J, Gurtman A, Lockhart S, et al. Phase I/II study of COVID-19 RNA vaccine BNT162b1 in adults. Nature. 2020;586: 589–593.
8. Venkatakrishnan AJ, Puranik A, Anand A, Zemmour D, Yao X, Wu X, et al. Knowledge synthesis of 100 million biomedical documents augments the deep expression profiling of coronavirus receptors. 2020 [cited 8 Dec 2020]. doi:10.7554/eLife.58040
9. Anand P, Puranik A, Aravamudan M, Venkatakrishnan AJ, Soundararajan V. SARS-CoV-2 strategically mimics proteolytic activation of human ENaC. 2020 [cited 8 Dec 2020]. doi:10.7554/eLife.58603
10. Ziegler CGK, Allon SJ, Nyquist SK, Mbano IM, Miao VN, Tzouanas CN, et al. SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. Cell. 2020;181. doi:10.1016/j.cell.2020.04.035
11. Zhao Y, Zhao Z, Wang Y, Zhou Y, Ma Y, Zuo W. Single-Cell RNA Expression Profiling of ACE2, the Receptor of SARS-CoV-2. Am J Respir Crit Care Med. 2020;202: 756.
12. Singh M, Bansal V, Feschotte C. A Single-Cell RNA Expression Map of Human Coronavirus Entry Factors. Cell Rep. 2020;32: 108175.
13. Yelin D, Wirtheim E, Vetter P, Kalil AC, Bruchfeld J, Runold M, et al. Long-term consequences of COVID-19: research needs. Lancet Infect Dis. 2020;20: 1115.
14. del Rio C, Collins LF, Malani P. Long-term Health Consequences of COVID-19. JAMA. 2020;324: 1723–1724.
15. Carfì A, Bernabei R, Landi F, for the Gemelli Against COVID-19 Post-Acute Care Study Group. Persistent Symptoms in Patients After Acute COVID-19. JAMA. 2020;324: 603–605.
16. Pawlowski C, Venkatakrishnan AJ, Ramudu E, Kirkup C, Puranik A, Kayal N, et al. Pre-existing conditions are associated with COVID patients' hospitalization, despite confirmed clearance of SARS-CoV-2 virus. medRxiv. 2020; 2020.10.28.20221655.
17. Pawlowski C, Venkatakrishnan AJ, Kirkup C, Berner G, Puranik A, O'Horo JC, et al. Enoxaparin is associated with lower rates of thrombosis, kidney injury, and mortality than Unfractionated Heparin in hospitalized COVID patients. medRxiv. 2020; 2020.10.06.20208025.
18. Pawlowski C, Puranik A, Bandi H, Venkatakrishnan AJ, Agarwal V, Kennedy R, et al.

Exploratory analysis of immunization records highlights decreased SARS-CoV-2 rates in individuals with recent non-COVID-19 vaccinations. medRxiv. 2020; 2020.07.27.20161976.

19. Awasthi S, Wagner T, Venkatakrishnan AJ, Puranik A, Hurchik M, Agarwal V, et al. Plasma IL-6 Levels following Corticosteroid Therapy as an Indicator of ICU Length of Stay in Critically ill COVID-19 Patients. medRxiv. 2020; 2020.07.02.20144733.
20. Agarwal V, Venkatakrishnan AJ, Puranik A, Lopez-Marquez A, Challener DW, Horo JC O, et al. Quantifying the prevalence of SARS-CoV-2 long-term shedding among non-hospitalized COVID-19 patients. medRxiv : the preprint server for health sciences. 2020 [cited 8 Dec 2020]. doi:10.1101/2020.06.02.20120774
21. Latz CA, DeCarlo C, Boitano L, Maximilian Png CY, Patell R, Conrad MF, et al. Blood type and outcomes in patients with COVID-19. Ann Hematol. 2020;99: 2113–2118.
22. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. Genomewide Association Study of Severe Covid-19 with Respiratory Failure. N Engl J Med. 2020;383. doi:10.1056/NEJMoa2020283
23. Zhao J, Yang Y, Huang H, Li D, Gu D, Lu X, et al. Relationship between the ABO Blood Group and the COVID-19 Susceptibility. medRxiv. 2020; 2020.03.11.20031096.
24. Ellinghaus D, Degenhardt F, Bujanda L, Buti M, Albillos A, Invernizzi P, et al. The ABO blood group locus and a chromosome 3 gene cluster associate with SARS-CoV-2 respiratory failure in an Italian-Spanish genome-wide association analysis. medRxiv. 2020; 2020.05.31.20114991.
25. Wagner T, Shweta FNU, Murugadoss K, Awasthi S, Venkatakrishnan AJ, Bade S, et al. Augmented curation of clinical notes from a massive EHR system reveals symptoms of impending COVID-19 diagnosis. 2020 [cited 8 Dec 2020]. doi:10.7554/eLife.58227
26. Pascarella G, Strumia A, Piliego C, Bruno F, Del Buono R, Costa F, et al. COVID-19 diagnosis and management: a comprehensive review. J Intern Med. 2020;288. doi:10.1111/joim.13091
27. Townsend L, Dyer AH, Jones K, Dunne J, Mooney A, Gaffney F, et al. Persistent fatigue following SARS-CoV-2 infection is common and independent of severity of initial infection. PLoS One. 2020;15: e0240784.
28. Hemoglobin test. 9 Oct 2019 [cited 8 Dec 2020]. Available: https://www.mayoclinic.org/tests-procedures/hemoglobin-test/about/pac-20385075
29. Hematocrit test. 12 Feb 2019 [cited 8 Dec 2020]. Available: https://www.mayoclinic.org/tests-procedures/hematocrit/about/pac-20384728
30. Khwaja A. KDIGO Clinical Practice Guidelines for Acute Kidney Injury. NEC. 2012;120: c179–c184.
31. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. Author Correction: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17: 352.