The COVID-19 immune landscape is dynamically and reversibly correlated with disease severity

Hamid Bolouri, … , Jane H. Buckner, the Benaroya Research Institute COVID-19 Research Team

*J Clin Invest.* 2021;131(3):e143648. [https://doi.org/10.1172/JCI143648](https://doi.org/10.1172/JCI143648).

**Graphical abstract**

Find the latest version:
[https://jci.me/143648/pdf](https://jci.me/143648/pdf)
The COVID-19 immune landscape is dynamically and reversibly correlated with disease severity

Hamid Bolouri,1,2 Cate Speake,3 David Skibinski,3 S. Alice Long,4 Anne M. Hocking,4 Daniel J. Campbell,5,6 Jessica A. Hamerman,5,6 Uma Malhotra,7,8 Jane H. Buckner,4,6 and the Benaroya Research Institute COVID-19 Research Team9

1Center for Systems Immunology, Benaroya Research Institute (BRI) at Virginia Mason, Seattle, Washington, USA. 2Division of Human Biology, Fred Hutchinson Cancer Research Center, Seattle, Washington, USA. 3Center for Interventional Immunology, 4Center for Translational Immunology, and 5Center for Fundamental Immunology, Benaroya Research Institute (BRI) at Virginia Mason, Seattle, Washington, USA. 6Department of Immunology, University of Washington School of Medicine, Seattle, Washington, USA. 7Department of Infectious Disease, Virginia Mason Medical Center, Seattle, Washington, USA. 8Department of Medicine, Section of Infectious Diseases, University of Washington, Seattle, Washington, USA. 9The BRI COVID-19 Research Team is detailed in the supplemental acknowledgments.

BACKGROUND. Despite a rapidly growing body of literature on coronavirus disease 2019 (COVID-19), our understanding of the immune correlates of disease severity, course, and outcome remains poor.

METHODS. Using mass cytometry, we assessed the immune landscape in longitudinal whole-blood specimens from 59 patients presenting with acute COVID-19 and classified based on maximal disease severity. Hospitalized patients negative for SARS-CoV-2 were used as controls.

RESULTS. We found that the immune landscape in COVID-19 formed 3 dominant clusters, which correlated with disease severity. Longitudinal analysis identified a pattern of productive innate and adaptive immune responses in individuals who had a moderate disease course, whereas those with severe disease had features suggestive of a protracted and dysregulated immune response. Further, we identified coordinate immune alterations accompanying clinical improvement and decline that were also seen in patients who received IL-6 pathway blockade.

CONCLUSION. The hospitalized COVID-19 negative cohort allowed us to identify immune alterations that were shared between severe COVID-19 and other critically ill patients. Collectively, our findings indicate that selection of immune interventions should be based in part on disease presentation and early disease trajectory due to the profound differences in the immune response in those with mild to moderate disease and those with the most severe disease.

FUNDING. Benaroya Family Foundation, the Leonard and Norma Klorfine Foundation, Glenn and Mary Lynn Mounger, and the National Institutes of Health.

Introduction

The coronavirus disease 2019 (COVID-19) pandemic has brought a worldwide focus not only on the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), but also on how immunity to the virus both promotes viral clearance and contributes to morbidity and mortality in infected individuals. There is a wide range of disease severity in SARS-CoV-2–infected individuals, ranging from asymptomatic infection to severe COVID-19 requiring mechanical ventilation, and in some cases, to death. Some factors have been identified that are associated with increased disease severity and poor outcome during COVID-19, including age, race, obesity, hypertension, and type 2 diabetes (1–11). However, we still do not understand the biologic factors that contribute to disease severity and outcome. It is becoming clear that not only does the severity of disease vary among SARS-CoV-2–infected individuals, but the immune response can also vary widely, leading to differing immune landscapes between patients. Therefore, it is important to understand how the immune landscape contributes to COVID-19 severity and outcome. Another important gap in our knowledge is how the immune landscape in COVID-19 resembles or is distinct from that seen in critically ill patients hospitalized for other reasons, since the immune landscape may change in the context of critical illness regardless of its etiology. In particular, it is important to determine if the early immune landscape can be used to inform which COVID-19 patients will have a severe disease course and would benefit from early interventions.

Although we can learn about immunity to SARS-CoV-2 by assessing a snapshot of the immune response at one point in time, the immune response to infection is dynamic and is best studied over time. Early immune responses to viruses are dominated by the innate immune system, including neutrophils,
monocytes, plasmacytoid dendritic cells (pDCs), and natural killer (NK) cells, whereas adaptive immune responses of T and B cells critical for viral clearance develop over days to weeks. Understanding how these populations change over time and relate to disease trajectory can give insight into the signature of a productive anti–SARS-CoV-2 immune response associated with clinical improvement, and whether immune dysregulation contributes to severe COVID-19. Additionally, early in the pandemic hospitalized patients were treated with a variety of experimental therapeutics, including the antiviral agent remdesivir, cytokine modulating therapies, and plasma from convalescent patients, all with varying efficacy in clinical studies and trials. However, how and if these treatments affect the immune landscape before and after therapeutic exposure has not been described.

To address these outstanding and important questions regarding the immune response during COVID-19, we used mass cytometry integrated with detailed clinical data to examine how the immune landscape changes over time in severe and moderate disease through natural progression and recovery, and also in the context of immune intervention.

Results

Patient demographics and clinical characteristics. We collected peripheral blood from 59 patients with COVID-19 (52 hospitalized patients and 7 ambulatory outpatients) at the Virginia Mason Medical Center, Seattle, Washington, USA, during the months of April and May 2020. Notably, we performed deep longitudinal sampling over the course of disease with an average of 4 time points per subject (range: 1–18; Figure 1), allowing for detailed immune trajectories of recovery. Patients were classified based on maximum disease severity using a 7-point ordinal scale (OS) representing the following outcomes: 1, not hospitalized with resumption of normal activities; 2, not hospitalized, but unable to resume normal activities; 3, hospitalized, not requiring supplemental oxygen; 4, hospitalized, requiring supplemental oxygen; 5, hospitalized, requiring nasal high-flow oxygen therapy, noninvasive mechanical ventilation, or both; 6, hospitalized, invasive mechanical ventilation; and 7, death (12).

Of the hospitalized patients, 24 were classified as having severe disease on the basis of requiring management in a critical care unit (CCU); all required mechanical ventilation (maximal OS ≥ 6), except one who was on high-flow oxygen (maximal OS = 5). The remaining 28 hospitalized patients were not in the CCU and were classified as having moderate COVID-19, with all requiring supplemental oxygen at some point in their hospital course (maximal OS = 3–5). The 7 ambulatory patients had mild disease (OS = 2) and did not require hospitalization. For a control group, we also collected blood from 17 hospitalized patients who tested negative for SARS-CoV-2; 4 of these patients were admitted to the CCU and the remainder to the floor. These patients were age- and sex-matched to the hospitalized COVID-19 groups and were admitted for a variety of conditions including respiratory (n = 4), cardiac (n = 4), gastrointestinal (n = 3), neurologic (n = 3), and miscellaneous conditions (n = 3).

The demographic and clinical characteristics of all the patient groups are summarized in Table 1. There was no significant difference in age or sex composition between severe, moderate, and mild COVID-19 groups. Regarding racial distribution, there was an overrepresentation in the severe COVID-19 group of African American (16.7%) and Hispanic (37.5%) individuals based on the Washington state population, which is 78.5% White, 4.4% African American, and 13% Hispanic (13). Duration of symptoms at time of presentation was longer in the severe disease group (median 9 days, range 3–22) compared with both the moderate (median 4 days, range 0–27) and mild (median 5 days, range 2–14) groups (P = 0.01). Duration of hospitalization was also significantly longer in the severe disease group (median 19 days, range 4–65) compared with the moderate disease group (median 6 days, range 2–28) (P < 0.01), although discharge was delayed for some patients due to restrictions placed on transfers to skilled nursing facility pending viral clearance from nasopharyngeal swabs.

Chronic medical conditions such as diabetes, hypertension, and cancer were common in the hospitalized COVID-19 cohorts. Diabetes was present in 45.8% of the severe group, 28.6% of the moderate group, and 28.6% of the mild group. Hypertension was present in 50% of the severe group and 67.9% of the moderate
Elevated white blood cell counts in COVID-19 correlate with increased neutrophils and decreased plasmacytoid dendritic cells and basophils. We assessed the immune landscape by combining clinical data with mass cytometry (CyTOF) performed on whole blood samples recovered from the clinical laboratory. The CyTOF panel was designed to assess the composition of the innate and lymphocyte compartments and determine the maturation, lineage, and activation status of these cell populations (Supplemental Table 1 and Supplemental Figures 2–4). To better understand the impact of disease, we performed a correlation analysis on the first sample collected for each patient in the COVID-19 cohort (n = 59; Figure 2 and Supplemental Figure 5). The heatmap in Figure 2A shows all significant correlations between clinical data (disease severity ordinal score, age, BMI, and CBC) and CyTOF immune cell percentages of the total CD45+ (pan-leukocyte marker) cell compartment, whereas the correlation network in Figure 2B focuses only on correlations among major leukocyte populations identified by CyTOF. We found correlations consistent with the current literature. For example, WBC counts and neutrophil counts were significantly correlated (Figure 2A), not surprisingly given that neutrophils comprise a large proportion of WBC, and both are elevated in severe COVID-19 (14, 15). Neutrophils in both the CBC and CyTOF data sets also inversely correlated with proportions of lymphocytes and T cells (Figure 2, A and B) supporting previous reports that the neutrophil-to-lymphocyte ratio is increased in severe COVID-19 (15–18). In addition, both pDCs and basophils negatively correlated with neutrophils, positively correlated with T cells, and positively correlated with each other (Figure 3, A–E). Together, these findings for pDCs and basophils are consistent with recent studies reporting depletion of these cell types in acute COVID-19 (19, 20). Although our CyTOF panel had limited ability to distinguish T cell lineage, T follicular helper (Tfh) cells were assessed. Notably, unlike other T cell populations, the percentage of Tfh cells in the memory CD4+ compartment showed a positive correlation with neutrophils, although this did not reach statistical significance (Figure 3F). Taken together, these observations indicate that coordinate and counter-acting changes in neutrophils, lymphocytes, pDCs, and basophils contribute to the immune signature of COVID-19.

The immune landscape differentiates individuals based on disease severity. In order to understand whether the immune signature in COVID-19 differed by disease severity, we determined the correlation between cell frequency and ordinal score at the time of sam-

Table 1. Cohort demographics and clinical characteristics

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>Median (range)</td>
<td>Median (range)</td>
<td>Median (range)</td>
<td>Median (range)</td>
</tr>
<tr>
<td>Number of days hospitalized</td>
<td>19 (4–65)</td>
<td>6 (2–28)</td>
<td>NA</td>
<td>5 (1–34)</td>
</tr>
<tr>
<td>Days from symptom onset to admission</td>
<td>9 (3–22)</td>
<td>4 (0–27)</td>
<td>5 (2–14)</td>
<td>NA</td>
</tr>
<tr>
<td>Disease score at admission</td>
<td>6 (3–6)</td>
<td>4 (2–4)</td>
<td>2 (2)</td>
<td>4 (3–4)</td>
</tr>
<tr>
<td>BMI</td>
<td>30.1 (18.1–54.5)</td>
<td>29.0 (17.0–54.1)</td>
<td>25.5 (22.5–37.2)</td>
<td>25.7 (21.1–69)</td>
</tr>
<tr>
<td>Outcome</td>
<td>Discharged 18 (75)</td>
<td>28 (100)</td>
<td>NA</td>
<td>16 (94.1)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>Female 12 (50)</td>
<td>14 (50)</td>
<td>3 (42.9)</td>
<td>10 (58.8)</td>
</tr>
<tr>
<td>Preexisting comorbidities</td>
<td>Asian 1 (4.2)</td>
<td>3 (10.7)</td>
<td>3 (42.9)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>African American 4 (16.7)</td>
<td>5 (17.9)</td>
<td>1 (14.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian/Pacific Islander 0 (0)</td>
<td>1 (3.6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Native American 2 (8.3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>White 6 (25)</td>
<td>17 (60.7)</td>
<td>2 (28.6)</td>
<td>14 (73.7)</td>
</tr>
<tr>
<td></td>
<td>Unknown/Other 11 (45.8)</td>
<td>2 (71)</td>
<td>1 (14.3)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino 9 (37.5)</td>
<td>2 (71)</td>
<td>1 (14.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Exposure to experimental medicine</td>
<td>Hydroxychloroquine 7 (29.2)</td>
<td>2 (71)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Remdesivir 17 (70.8)</td>
<td>11 (39.3)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Tocilizumab 8 (33.3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Convalescent plasma 15 (62.5)</td>
<td>4 (14.3)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
</tbody>
</table>

Table 2. Clinical background and laboratory data

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Days from symptom onset to admission</td>
<td>9 (3–22)</td>
<td>4 (0–27)</td>
<td>5 (2–14)</td>
<td>NA</td>
</tr>
<tr>
<td>Disease score at admission</td>
<td>6 (3–6)</td>
<td>4 (2–4)</td>
<td>2 (2)</td>
<td>4 (3–4)</td>
</tr>
<tr>
<td>BMI</td>
<td>30.1 (18.1–54.5)</td>
<td>29.0 (17.0–54.1)</td>
<td>25.5 (22.5–37.2)</td>
<td>25.7 (21.1–69)</td>
</tr>
<tr>
<td>Outcome</td>
<td>Discharged 18 (75)</td>
<td>28 (100)</td>
<td>NA</td>
<td>16 (94.1)</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td>Female 12 (50)</td>
<td>14 (50)</td>
<td>3 (42.9)</td>
<td>10 (58.8)</td>
</tr>
<tr>
<td>Preexisting comorbidities</td>
<td>Asian 1 (4.2)</td>
<td>3 (10.7)</td>
<td>3 (42.9)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>African American 4 (16.7)</td>
<td>5 (17.9)</td>
<td>1 (14.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Native Hawaiian/Pacific Islander 0 (0)</td>
<td>1 (3.6)</td>
<td>0 (0)</td>
<td>0 (0)</td>
</tr>
<tr>
<td></td>
<td>Native American 2 (8.3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>White 6 (25)</td>
<td>17 (60.7)</td>
<td>2 (28.6)</td>
<td>14 (73.7)</td>
</tr>
<tr>
<td></td>
<td>Unknown/Other 11 (45.8)</td>
<td>2 (71)</td>
<td>1 (14.3)</td>
<td>1 (5.9)</td>
</tr>
<tr>
<td></td>
<td>Hispanic/Latino 9 (37.5)</td>
<td>2 (71)</td>
<td>1 (14.3)</td>
<td>0 (0)</td>
</tr>
<tr>
<td>Exposure to experimental medicine</td>
<td>Hydroxychloroquine 7 (29.2)</td>
<td>2 (71)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Remdesivir 17 (70.8)</td>
<td>11 (39.3)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Tocilizumab 8 (33.3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>Convalescent plasma 15 (62.5)</td>
<td>4 (14.3)</td>
<td>0 (0)</td>
<td>NA</td>
</tr>
</tbody>
</table>
plunging. Increasing neutrophil frequency was positively correlated with increasing disease severity (Pearson correlation $r = -0.46$, FDR-adjusted $P < 0.01$), whereas T cells, NK cells, pDCs, and basophils were lower in severe disease (all FDR-adjusted $P$ values < 0.005, Figure 4). To determine if the immune landscape early in disease distinguishes severe from mild disease, we next performed a cross-sectional analysis of our population categorized based on an individual's highest disease score during the course of their illness using data from the first sample collected for each patient (Figure 5, Supplemental Figure 5). The CBC data showed the greatest difference with disease severity in WBC counts with an increase in the absolute neutrophils and monocyte counts and low absolute lymphocyte counts (Figure 5A). However, these CBC results frequently fell within the normal range and notably, the hospitalized COVID-19–negative population showed very similar changes to those seen with severe COVID-19, suggesting that these findings are not unique to COVID-19 but are instead reflective of critical illness.

In contrast, the cross-sectional analysis of the CyTOF data set identified 2 different patterns of immune alterations in the COVID-19 cohort: those that were also present in the hospitalized COVID-19–negative cohort and those that were unique to severe COVID-19. Immune cell populations that were similar between
The Journal of Clinical Investigation

CLINICAL MEDICINE

CLINICAL MEDICINE

5

J Clin Invest. 2021;131(3):e143648  https://doi.org/10.1172/JCI143648

erally differentiated individuals based on their disease severity, with more moderate disease courses and good outcome associated with clusters A and B, while those with the most severe disease and death were associated with cluster C. These findings indicate that there is not one single immune signature in COVID-19, but that the immune response differs in individuals based on the ultimate disease severity.

Immune trajectories discriminate moderate and severe COVID-19. To better understand the kinetics and coordinated changes in immune signatures, we tracked immune cell types in the blood over time based on date of admittance to the hospital. We focused on exploring differences in longitudinal analysis of moderate and severe patients based on distinct clustering between these groups as shown in Figure 6. The time course was limited to 15 days after admittance for sufficient and comparable sampling in both the moderate and severe cohorts (Supplemental Figure 6).

Key to understanding features that distinguish moderate from severe COVID-19 is an appreciation of the evolution of the immune response over time, as shown in a uniform manifold approximation and projection (UMAP) visualization of immune changes with disease severity within cell types of an individual patient (Figure 7). Using gated data from Figures 2–7, we focused on specific cell types and markers of innate and adaptive immunity. Figure 8 shows Loess-smoothed trajectories whereas Supplemental Figures 7–9 show individual and averaged plots. We found that patients with moderate COVID-19 had a dynamic immune response that resolved over time, typical of a productive antiviral response, whereas patients with severe COVID-19 had an aberrant immune response, diverging early from that seen in moderate COVID-19 subjects and continuing to diverge beyond the first 15 days of hospitalization. Specifically in the moderate COVID-19 cohort, there was an early reduction in circulating neutrophils with a concomitant increase in circulating monocytes, total DCs, and basophils, with maximal change at 4–5 days after hospitalization (Figure 8A and Supplemental Figures 7–9). In addition, NK cells increased following these early myeloid cell changes, peaking at 5–6 days after hospitalization (Figure 8A and Supplemental Figures 8 and 9). In contrast with the severe COVID-19 cohort, these innate cell populations were less dynamic with little variation during the first 15 days of hospitalization (Figure 8A and Supplemental Figures 8 and 9).

Unsupervised hierarchical clustering of the CyTOF data for each subject’s initial sample identified 3 major clusters of patients (Figure 6): a T cell predominant cluster with a relative decrease in neutrophils (cluster A), a cluster with mixed features including a predominance of monocyte, DC, and NK cells (cluster B), and a third cluster whose patients had high levels of neutrophils and a relative paucity of other cell types (cluster C). These clusters generally differentiated individuals based on their disease severity, with more moderate disease courses and good outcome associated with clusters A and B, while those with the most severe disease and death were associated with cluster C. These findings indicate that there is not one single immune signature in COVID-19, but that the immune response differs in individuals based on the ultimate disease severity.

Figure 3. Correlations among immune cell populations in patients with COVID-19 demonstrate a relationship between disease severity and an increase in neutrophils and a depletion of pDCs and basophils. (A–F) Plots display FDR-adjusted Pearson correlations and linear regression lines with 95% confidence interval shading. Data points are colored according to the ordinal score observed for each patient at admission.
immune response in their blood, peaking early after hospitalization and then resolving to levels similar to those seen in patients with moderate COVID-19 by day 15 after hospitalization (Figure 8B). Thus overall, patients with moderate COVID-19 showed signatures of a productive innate response that resolved within the 2 weeks of hospitalization, whereas patients with severe disease showed signs of an aberrant response after hospital admittance that persisted for at least the first 2 weeks in hospital.

**Immune signatures of clinical improvement in patients with COVID-19.** To identify key immune cell populations that are associated with either clinical improvement or decline, we focused our analysis on samples taken from individuals before and after a change in ordinal score, reflective of disease severity. We assessed changes in the absolute abundance of immune cell populations by CBC or in the frequency of immune cell subsets in our CyTOF analyses across these key clinical times. We identified subjects that had samples drawn across a score improvement of 2 or more, or a score decline of 1 or more (Figure 9A). This analysis identified several populations whose abundance or frequency was significantly altered upon changes in ordinal score (Figure 9B). Consistent with the lymphopenia observed in severe COVID-19, we found that absolute lymphocytes decreased with clinical decline whereas an increase in the absolute number of lymphocytes was associated with clinical improvement. The increase in lymphocytes was mediated by a general increase in the frequency of naive and memory CD4⁺ T cells as well as NK cells, but not B cells. The frequency of pDCs also increased in subjects around the time of clinical improvement, whereas an increase in the frequency of neutrophils was observed during recovery from severe COVID-19 (Figure 9C). This analysis demonstrates that the immune landscape is dynamic in COVID-19, and that resolution of key features of severe disease is coincident with improvement in clinical status.

**Early immune signatures of tocilizumab, but not convalescent plasma, treatment in patients with severe COVID-19.** To determine if there were immune signatures of tocilizumab or convalescent plasma treatment, we identified 7 patients treated with tocilizumab and 7 patients treated with convalescent plasma in our cohort and then resolving to levels similar to those seen in patients with moderate COVID-19 by day 15 after hospitalization (Figure 8A and Supplemental Figures 8 and 9). Thus overall, patients with moderate COVID-19 showed a signature of a productive innate immune response in their blood, peaking early after hospitalization, whereas patients with severe disease showed a blunted and delayed innate response.

Adaptive responses also differed between moderate and severe COVID-19 over time (Figure 8B and Supplemental Figures 8 and 9). In patients with moderate COVID-19, total T cells expanded and contracted, consistent with an expected antiviral T cell response, with a later enrichment of memory CD4⁺ T cells (Figure 8B). Memory B cells increased more robustly over time in the moderate COVID-19 cohort throughout hospitalization, suggestive of sustained interaction with memory CD4⁺ T cells and antibody production. In contrast, patients with severe COVID-19 consistently had lower levels of both T cells and memory B cells over the course of hospitalization, suggesting a diminished or delayed adaptive immune response to the virus. The Th1 response in the severe COVID-19 cohort was greater than that of the moderate COVID-19 cohort at all time points, perhaps indicating unresolved T cell help or Th1 cells sustained by high IL-6 in critically ill patients. In addition, in the severe COVID-19 cohort, Treg cells as a percentage of total CD4⁺ T cells were increased over time as compared with the moderate COVID-19 cohort (Figure 8B), likely in response to ongoing inflammation due to viral persistence.

Consistent with this idea, the percentage of CD8⁺ T cells expressing HLA-DR, a marker of activation, also increased over time in the severe COVID-19 cohort (Figure 8B), as did CD8⁺ T cells expressing CD38 and PD-1 (Supplemental Figure 7B), whereas total memory CD8⁺ T cell increases were similar between moderate and severe patients (Figure 8B). Overall, our longitudinal analysis revealed that the immune trajectory differs between moderate and severe patients during the first 2 weeks after initial hospitalization. Patients with moderate disease showed signatures of a productive antiviral response that resolved within the 2 weeks of the study time, whereas patients with severe disease showed signs of an aberrant response after hospital admittance that persisted for at least the first 2 weeks in hospital.

In 274 samples from 59 patients with COVID-19, the abundances of (A) neutrophils, (B) T cells, (C) NK cells, (D) pDCs, and (E) basophils are highly correlated with disease severity (all P values FDR adjusted). Red plot points mark values for samples further analyzed in improving versus declining patients (Figure 9).

**Figure 4. Cross-sectional immune correlates of COVID-19 disease severity.** In 274 samples from 59 patients with COVID-19, the abundances of (A) neutrophils, (B) T cells, (C) NK cells, (D) pDCs, and (E) basophils are highly correlated with disease severity (all P values FDR adjusted). Red plot points mark values for samples further analyzed in improving versus declining patients (Figure 9).
who had CyTOF samples both before and after treatment (Supplemental Tables 2 and 3). Notably, these patients all had severe disease and there were stringent criteria for the use of tocilizumab, including rapidly escalating oxygen needs combined with an IL-6 level greater than 20 times the upper limit of normal (ULN); and C-reactive protein (CRP) greater than 125 mg/L (ULN = 7.9). Marked elevations in ferritin, LDH, and D-dimer were also weighted in the decision-making process. All patients were also treated with remdesivir, with the exception of one patient in the convalescent plasma group. Additionally, 6 of 7 patients in the tocilizumab group analyzed were treated with convalescent plasma prior to tocilizumab treatment (1–6 days before tocilizumab). None of the 7 patients in the convalescent plasma group were treated with tocilizumab during the time points analyzed. Also important, patient care was similar between the 2 groups as the use of tocilizumab in our hospital does not result in alterations to patient care.

We first assessed serum CRP levels in these 2 groups as a measure of the effectiveness of tocilizumab treatment, which should reduce this marker of systemic inflammation. Indeed, treatment with tocilizumab swiftly reduced serum CRP in all patients (Figure 6A). Serum ferritin was also reduced mainly...
us to assess the fold change in response to treatment for each patient. In the tocilizumab group there were several populations of immune cells that differed significantly before and after treatment (Figure 10B). In contrast, there were no significant changes after convalescent plasma treatment in the immune cell populations analyzed by CyTOF (Figure 10C and Supplemental Figure 10, B and C). The significant changes in response to tocilizumab treatment included a reduction in the percentage of neutrophils and an increase in the percentage of total T cells, eosinophils, basophils, and DCs among CD45+ cells (Figure 10, B, D, and E). There were also increases in several CD4+ and CD8+ T cell sub-populations, and no changes in any B cell populations after tocilizumab (Figure 10B). Moreover, our findings for T cells, B cells, neutrophils, and basophils were consistent with our signature of clinical improvement (Figure 9). However, there was not complete overlap between the tocilizumab signature and the clinical improvement signature, as NK cells and pDCs were not significantly changed by tocilizumab (Figure 10B, data not shown) but were increased in improving patients (Figure 10B). In the tocilizumab group, we also identified increased populations associated with T cell activation, including HLA-DR+ and CD38+ CD4+ and CD8+ T cells (Figure 10, B, D, and E). In summary, we observed a clear acute signature of tocilizumab treatment that shares some but not all features of the immunologic changes seen with clinical improvement, whereas there is no acute change in the immune landscape with convalescent plasma treatment, in patients with severe COVID-19.

Discussion

A growing body of literature indicates that the immune landscape is profoundly altered by COVID-19 and differs between individuals dependent on disease severity (18, 20, 25, 26). Whether the immune landscape is a reflection of disease severity, a source of severe disease, or a combination of the 2 is still not fully understood. Here, we utilized recovered samples from the clinical laboratory to rapidly assess peripheral blood cell populations by CyTOF and these data were analyzed in conjunction with clinical laboratories and disease severity scores. Importantly, we were able to collect samples longitudinally among the hospitalized individuals, allowing us to examine the evolution of immune responses through natural progression and recovery, and in the context of immune intervention. Novel aspects of our study included: (a) deep longitudinal sampling allowing for detailed immune trajectories of recovery, (b) a control cohort of moderate and severely ill hospitalized COVID-19 negative patients, and (c) analysis of immune signatures associated with tocilizumab and convalescent plasma treatments.

Notably, we found that at the time of initial sampling the immune landscape in COVID-19 forms 3 dominant clusters that relate to disease severity. When we examined individual cell populations based on disease severity, we found, as others have, that the neutrophil-to-lymphocyte ratio is increased in individuals with severe COVID-19 (15–18). Furthermore, this inverse relationship with neutrophils applies to basophils, DC, NK cells, and monocytes, and only modestly with B cells, and is most pronounced among T lymphocytes with the exception of Tfh cells, which are positively correlated to neutrophil numbers, a finding
Our longitudinal assessment further allowed us to identify patterns that distinguished severe and moderate disease. Individuals with a moderate disease course showed a pattern consistent with productive innate and adaptive immune response characterized by early and transient increases in monocytes and NK cells with later sustained increases in memory T and B cells. Those with severe disease have features suggestive of a dysregulated immune response characterized by delayed and prolonged increases in Tfh cells, HLA-DRlo monocytes, and activated CD8+ T cells. Although the time from first symptom and first sample was delayed in severe patients as compared with moderate patients, this was not reflected in a simple shift of the immune trajectory. Instead, changes in multiple cell types of moderate subjects were transitory while evidence of persistent activation in different immune cells were progressive and unresolved in severe patients. This suggests that the degree of inflammation or persistence of virus markedly changed the immune landscape over time in severe as compared with moderate disease. Importantly, the persistent features of severe disease are reversed with improvements in clinical score and can be modulated in part with immune interventions, such as IL-6 pathway blockade.

Our findings for the tocilizumab study were intriguing, especially in the light of the recent disappointing results from the first randomized double blind phase 3 trial (27). Our data suggest that changes in the immune landscape after tocilizumab treatment, with the exception of NK cell and pDC recovery, are consistent with the immune signature of clinical improvement that we identified. Of note, only 4 of 7 of these tocilizumab-treated patients improved clinically in the acute time frame of our analysis (within 9 days of treatment) with the remaining 3 patients only showing clinical improvement at later times. This suggests that although tocilizumab treatment induces an acute signature of clinical improvement in both serum CRP and specific immune cell populations, this signature is disconnected from immediate clinical change, indicating that the immune changes with tocilizumab may be inadequate to support full recovery. Caveats for our tocilizumab analysis include the small cohort size, and that all but one of these patients were treated with convalescent plasma prior to tocilizumab treatment. Therefore, it is possible that convalescent plasma acts synergistically with tocilizumab to cause the immune signature we identified. Interestingly and in contrast to tocilizumab, we saw no clear immune signature of convalescent plasma within 7 days, suggesting either our cohort was too small to see changes, the immune populations change after the times we analyzed, or convalescent plasma does not act at the level of blood leukocyte populations. It is clear that further investigation is needed to determine whether tocilizumab has a therapeutic role in COVID-19, and in what patient population it would be useful, and this may be determined in part by the character and trajectory of the immune landscape of the patient.

The demographics of our patients with COVID-19 were consistent with published case reports. African Americans and Hispanics were overrepresented in the severe COVID-19 group relative to the population of Washington state, which is consistent with reports from other states in the USA (10, 11). We also found that type 2 diabetes was more common in those with severe disease compared with those with moderate or mild disease. Notably, all groups have
higher diabetes prevalence than the US or Washington rates (28); the highest prevalence in Washington state is among 65–74 year olds at 21.5%, which is more than doubled in the cohort with severe disease described here. Diabetes and obesity have consistently been identified as risk factors for COVID-19 severity (4–9); reduced T cell function and chronic inflammation have been postulated as potential mechanisms driving this increased risk (29). In addition, some glucose-lowering agents used in diabetes are known to impact the immune system (reviewed in ref. 30). Full analysis of the differential impact of diabetes and its treatment on our immune signatures is beyond the scope of this work but merits further analysis.

There are limitations to this study. Due to the urgency of the pandemic, we chose to use recovered clinical samples for our study and thus the collection schedule and sample availability was dictated by the treatment needs of the patient. This meant that we did not have the same time points for every patient, and that we could not match between groups the medications that individuals were already taking due to preexisting comorbidities, some of which may impact the immune responses seen here. In addition, the differences between the mild, moderate, and severe COVID-19 groups may reflect the time from disease onset, which significantly varied between these groups, and/or differences in viral burden, which we could not assess. The subjects in our hospitalized control group were not matched to the SARS-CoV-2–positive groups by race, although they are well matched by age.

In summary, we have identified unique features of the immune landscape in moderate versus severe COVID-19 along with features that are common to moderate and severe non–COVID-19 illness. Importantly, our findings indicate that selection of immune interventions should be based in part on disease presentation and early disease trajectory due to the profound differences in the immune response in those with mild to moderate disease and those with the most severe disease. Finally, our characterization of the variety of immune signatures in COVID-19 provides insight into the types of immune interventions that may be beneficial in the treatment of severe disease.
collected starting as soon as possible after hospital admission, then if feasible, daily for the first week, and then at 3–4 day intervals (Figure 1). A single sample was obtained at time of first outpatient visit for the ambulatory COVID-19 subjects. A maximum of 2 samples was obtained from the hospitalized COVID-19–negative control subjects. All assays were run and analyzed in a blinded manner.

**CyTOF staining, acquisition, and subset identification.** Peripheral blood was collected from each donor into sterile vacutainer tubes containing the anticoagulant EDTA. Blood cells were washed twice with PBS and stained for viability exclusion with a 100 μM cisplatin solution (Enzo Life Sciences) for 1 minute at room temperature. Cisplatin was quenched with 5 volumes MaxPar Cell Staining Buffer (CSB; Fluidigm),

**Methods**

**Study design.** Using our newly developed 33-parameter CyTOF panel, we characterized the immune response longitudinally in 59 adults with acute COVID-19, including 24 hospitalized patients with severe disease, 28 hospitalized patients with moderate disease, and 7 ambulatory patients with mild disease not requiring hospitalization. All COVID-19 subjects were positive for SARS-CoV-2 and our control cohort of 17 hospitalized patients tested negative for SARS-CoV-2. Healthy control subjects were age- and sex-matched to the hospitalized COVID-19 subjects. Importantly, the samples used were collected prior to the start of the COVID-19 pandemic in December 2019. For the hospitalized COVID-19 cohort, longitudinal samples were

---

**Figure 9. Immune signatures of clinical decline and improvement.** (A) Schematic outlining approach for identifying immune signatures of clinical decline and improvement focusing on changes in monitored parameters in longitudinal samples taken before and after changes in clinical score. (B) Log₂-fold change in the indicated cell populations as measured by CyTOF or CBC analysis in longitudinal samples taken before and after improving (green) or declining (red) clinical scores. Asterisks indicate a significant difference in the fold changes (2-tailed, unpaired Wilcoxon rank sum FDR-adjusted P < 0.05) between improving (n = 7) and declining (n = 10) patient groups for the indicated cell populations. (C) Longitudinal analyses of the frequency of neutrophils, T cells, NK cells, and pDCs vs. clinical score in 3 individual patients. Black line shows disease score (left y axis) and red line shows immune cell frequency (right y axis).
and the cells then stained with a titered, aliquoted, and frozen cocktail of monoclonal antibodies conjugated to metal isotopes for 20 minutes at 4°C. Red blood cell lysis was performed using RBC Lysis/Fixation solution (BioLegend) for 5 minutes at room temperature followed by a wash with CSB. The resulting leukocytes were fixed overnight at 4°C with MaxPar Fix and Perm Solution (Fluidigm) containing 125 nM Cell-ID Intercalator-Ir (Fluidigm). Following fixation, cells were washed with CSB, resuspended in milli-Q water, and stored at 4°C until

Figure 10. Early immune responses to tocilizumab but not convalescent plasma in severe COVID-19. (A) Serum CRP in patients receiving tocilizumab (top; \( n = 7 \)) or convalescent plasma (bottom; \( n = 7 \)) measured in clinical labs relative to day of treatment. Each line represents an individual patient. (B and C) Change in blood immune populations measured by CyTOF after treatment with tocilizumab (B) or convalescent plasma (C). The fold change in each population for each subject was determined by dividing the percentage of each population in the first posttreatment sample at day +2 or more after treatment with the closest pretreatment sample available as detailed in Supplemental Table 3. All are shown as percentage of CD45+ cells unless otherwise indicated. (D) Plots showing the percent of the indicated populations in tocilizumab-treated patients before and after treatment, using the time points used for analysis in B and Supplemental Table 3. (E) Plots showing all the data points available for tocilizumab-treated patients for the indicated populations shown in D and Supplemental Table 3. Each line represents an individual patient and the color of the line reflects the clinical ordinal score at the time of sampling. *\( P < 0.05 \) Wilcoxon matched pairs test, adjusted for multiple comparisons.
acquisition. All antibodies from BioLegend and BD Biosciences were conjugated to their respective metal isotopes using the MaxPar X8 Multimetal Labeling Kit (Fluidigm). Samples were stained within 48 hours of blood draw. Sample stability with the CyTOF assay was established with 3 COVID-19 samples assayed on the day of collection (baseline, day 0), one day after collection (day 1) or 2 days after collection (day 2). Populations and markers on populations were standard and based on Staser et al. (31). All populations and markers with a frequency greater than 5% of live CD45+ cells had a CV less than 40% between baseline and each time point. This variation was less than that of biological comparisons. MaxPar Four Element Calibration Beads (Fluidigm) were added to each sample immediately before acquisition. All samples were acquired on a Helios CyTOF mass cytometer (Fluidigm) with a target cell acquisition of 100,000 live events at a rate of 500 events/second to capture greater than 50 cells per gated population or marker. The CyTOF panel is shown in Supplemental Table 1 and gating strategies are shown in Supplemental Figures 2–4. To determine gates for activation markers such as CD25, CD69, and CD38 on T cells and PD-L1 on myeloid cells, we first analyzed 12 samples from 6 subjects with moderate COVID-19 and 6 subjects with severe COVID-19. Gates were set based on a comparison between samples that were clearly highly activated and those that were clearly nonactivated. These gates were then applied to all samples in the study and used consistently for all populations analyzed. Specifically, gates for CD25, CD38, CD69, HLA-DR, PD-1, and PD-L1 were the same for all cell types where they were applied. For example, the CD38 gate was the same for CD4+ T cells, CD8+ T cells and Tfh cells, the CD25 gate was the same for CD4+ and CD8+ T cells, and the CD69 gate was the same for CD4+ T cells, CD8+ T cells, NK cells, eosinophils, neutrophils, etc. Data were analyzed using a FlowJo software versions 10.6.0 and 10.6.1 (FlowJo).

Statistics. Apart from the paired-sample tests in Figure 10, all P values were calculated using unpaired, 2-tailed Wilcoxon rank sum tests. In all cases, corrections for multiple testing were performed using the FDR method. For between group comparisons of the clinical data, P values were calculated using the Kruskal-Wallis 1-way analysis of variance test. The correlation graph in Figure 2A was built from the matrix of Pearson correlations in Figure 2B using the R igraph package (32). The heatmap in Figure 6 was generated using Euclidean distance and the clustering method Ward.D2.

Time-series data from each patient were organized in terms of the relative number of days from the date of the first sample (hereon denoted pseudo-time), and then aligned by first sample. To reduce the potential effects of outlier samples, median values were calculated for each severity category and each day for the samples available. If no samples were available at a given pseudo-time point, we inferred a value using linear interpolation between the before and after pseudo-time points. The vertical bars at each pseudo-time point are equal to 1 standard deviation from the indicated median value. Plot points with no error bars are those with only 1 sample or represent an inferred value. Loess smoothing was performed on the median values for each disease severity class using the geom_smooth function in the R ggplot library (33).

The UMAP plots in Figure 7 were generated directly from the CyTOF signal intensities following archsinh transformation with a cofactor value of 5. To ensure against batch and other potential confounding effects, we specifically selected samples collected and stained in a highly uniform fashion from a single donor and z score normalized probe intensities for each sample prior to UMAP projection to 2D.

Study approval. Samples from COVID-19 subjects and from hospitalized COVID-19-negative control subjects were recovered from the Virginia Mason Medical Center Central Processing Lab after all tests required for clinical care were complete, under approval by the Benaroya Research Institute (BRI) protocol IRB20-036. All healthy control samples were from healthy subjects in the BRI Immune-Mediated Disease Registry and Repository who had given written informed consent in accordance with the Declaration of Helsinki and according to the BRI IRB-approved protocol IRB07109.

Author contributions

SAL, CS, UM, JAH, DJC, and JHB conceptualized and designed the study. UM and JHB were responsible for subject selection and clinical interpretation. The BRI COVID-19 Research Team obtained IRB approval, collected samples and clinical data, conducted the mass cytometry, and integrated the clinical data with the CyTOF data. SAL, DS, and JAH performed the gating for the immune cell populations measured by CyTOF. HB, SAL, DS, and JAH performed the computational and statistical analysis. JAH, AMH, and JHB wrote the manuscript with assistance from all coauthors. JHB obtained funding and supervised the study.

Acknowledgments

We would like to acknowledge the Benaroya Family Foundation, the Leonard and Norma Klorfine Foundation, and Glenn and Mary Lynn Mounger for their funding of this project. This work was supported by NIH grants R01AI150178-01S1 (to JAH) and U19AI125378-05S1 (to DJC and JAH). We acknowledge the Allen Institute for Immunology for their funding support for the development of the CyTOF panel used in this study. We also would like to thank Carmen Mikacenic for help with obtaining IRB approval, and Henry T. Bahnson for help with statistical analysis. See Supplemental Acknowledgments for the BRI COVID-19 Research Team details. The graphical abstract was created with BioRender.com.

Address correspondence to: Jane H. Buckner, Benaroya Research Institute at Virginia Mason, 1201 Ninth Avenue, Seattle, WA 98101. Phone: 206.287.1033; Email: jhbuckner@benaroyaresearch.org.