

## CORONAVIRUS

# Immunophenotyping assessment in a COVID-19 cohort (IMPACC): A prospective longitudinal study

IMPACC Manuscript Writing Team<sup>†</sup> on behalf of the IMPACC Network Steering Committee\*<sup>‡</sup>

The Immunophenotyping Assessment in a COVID-19 Cohort (IMPACC) is a prospective longitudinal study designed to enroll 1000 hospitalized patients with COVID-19 (NCT04378777). IMPACC collects detailed clinical, laboratory, and radiographic data along with longitudinal biologic sampling of blood and respiratory secretions for in-depth testing. Clinical and laboratory data are integrated to identify immunologic, virologic, proteomic, metabolomic, and genomic features of COVID-19–related susceptibility, severity, and disease progression. The goals of IMPACC are to better understand the contributions of pathogen dynamics and host immune responses to the severity and course of COVID-19 and to generate hypotheses for identification of biomarkers and effective therapeutics, including optimal timing of such interventions. In this report, we summarize the IMPACC study design and protocols including clinical criteria and recruitment, multisite standardized sample collection and processing, virologic and immunologic assays, harmonization of assay protocols, high-level analyses, and the data sharing plans.

## INTRODUCTION

The coronavirus disease 2019 (COVID-19) pandemic urgently demands comprehensive knowledge about the immunology, virology, and genetics of this disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Mobilization of the global scientific community has produced substantial translational findings with unprecedented speed but often from limited patient populations (1–5). The National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), launched a prospective longitudinal cohort study [Immunophenotyping Assessment in a COVID-19 Cohort (IMPACC)] in May 2020. IMPACC aims to enroll at least 1000 adults hospitalized for known or presumptive COVID-19 in about 20 hospitals associated with 15 U.S. biomedical research centers and collect clinical data and biological samples for up to 12 months after discharge. Harmonized clinical data are obtained and biologic samples are assayed at 11 centralized Core immunoassay laboratories (Fig. 1). The goal of the study is to better understand the contributions of the pathogen and host immune response in modulating the manifestations, severity, and course of COVID-19 and to identify potential biomarkers as well as inform therapeutic interventions. In this report, we summarize the overall study design, multicenter coordination and harmonization of clinical data and biologic sample collection and processing, protocols for virologic and immunologic core assays, and approaches for a high-level integrated analysis plan.

## STUDY DESIGN

### Study overview and rationale

IMPACC is an observational cohort study designed to survey clinical and immunologic manifestations of COVID-19 in hospitalized patients (NCT04378777), collaboratively developed by the NIAID and investigators from the Human Immunology Project Consortium

\*Corresponding author. Email: augustine@niaid.nih.gov (A.D.A.); patrice.becker@nih.gov (P.M.B.)

†The IMPACC Manuscript Writing Team members and affiliations are located at the end of this paper. Individual contributions are listed in the Acknowledgments.

‡The IMPACC Network Steering Committee members and affiliations are located at the end of this paper.

(HIPC), Asthma and Allergic Diseases Cooperative Research Centers (AADCRC), and other NIAID-funded programs. The IMPACC network brings together expert clinicians, geneticists, and immunologists to assess the relationship between the clinical course and immune response to SARS-CoV-2 in racially, ethnically, and geographically diverse adult patient populations across the United States. IMPACC's primary objectives are to (i) describe the relationship between specific immunologic assessments and severity of illness in hospitalized patients with COVID-19, controlling for time of illness onset, and concurrent participation in clinical trials or off-label use of investigational (or approved) therapeutic agents for COVID-19 and (ii) describe the relationship between burden of disease, assessed by duration of virus shedding in nasal secretions, and severity of illness in hospitalized patients with COVID-19. The study incorporates clinical data collection elements harmonized with the publicly available International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) case report forms (CRFs), as well as standardized biologic sample collection and processing protocols, to minimize confounding variables across sites. Local upper and lower airways and systemic immunologic parameters are surveyed via comprehensive, unbiased sample-sparing omics assessments (Supplementary Methods). An integrative data analysis plan is being developed to map immunologic endotypes (i.e., distinct functional subtypes of human immune responses) to clinical phenotypes in this patient cohort. In addition to focusing on unique features defining acute disease course, the cohort aims to follow participants for up to a year after hospital discharge to assess measures of both functional and immunologic recovery.

### Clinical study design

The high-level study design is shown in Fig. 2. Participants are enrolled within 48 hours of hospital admission. Demographic information, COVID-19 symptoms and onset, and detailed medical history including comorbidities are collected from the medical record and/or patient interviews for all participants at baseline. Only cases with confirmed positive SARS-CoV-2 polymerase chain reaction (PCR) are followed longitudinally. Participants undergo extensive serial assessments to capture clinical data as shown in Box 1 (including clinical laboratory values, radiographic findings, medication

Copyright © 2021  
The Authors, some  
rights reserved;  
exclusive licensee  
American Association  
for the Advancement  
of Science. No claim to  
original U.S. Government  
Works. Distributed  
under a Creative  
Commons Attribution  
License 4.0 (CC BY).

use, oxygen and ventilatory support requirements, and complications) and biologic samples [blood, mid-turbinate nasal swabs, and, for intubated patients, endotracheal aspirates (EAs)]. Clinical data and samples are collected at enrollment and at days 4, 7, 14, 21, and 28 while participants are hospitalized. If a participant requires an escalation to intensive care unit (ICU)-level care or is discharged and readmitted to the hospital >48 hours after discharge, additional samples are collected within 24 and 96 hours of care escalation or readmission. Key clinical outcome data collected during hospitalization include mortality, level of care (floor, ICU), respiratory support requirements, extrapulmonary organ dysfunction, length of stay, and patient status based on an ordinal scale (6). If a participant is discharged from the hospital before day 14 or 28, attempts are made to collect additional data and samples at days 14 and 28 on an outpatient basis. Convalescent questionnaires and biologic samples are collected at 3-month intervals up to 12 months after hospital discharge.

For outpatient visits, both symptoms and functional recovery are surveyed using patient-reported outcome measures collected at the study site, by telephone, or via electronic data capture using an app or web portal. Specific patient-reported outcome measures to be assessed at outpatient follow-up include presence of COVID-19 symptoms, health-related quality of life using the EQ-5D-5L (7), and several PROMIS (8) surveys to capture physical, emotional, cognitive, psychosocial, and respiratory functional status. Efforts are made across sites to approach as many as possible hospitalized patients with known or confirmed COVID-19. Study information is provided to non-English-speaking patients in their native language.

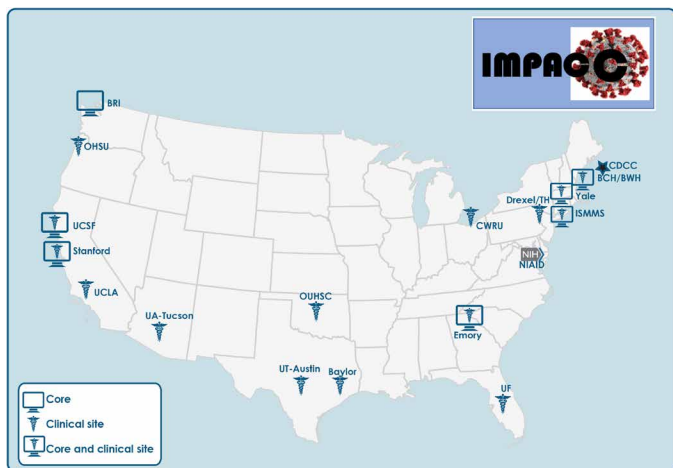
NIAID staff conferred with the Department of Health and Human Services Office for Human Research Protections (OHRP) regarding potential applicability of the public health surveillance exception [45CFR46.102(l)(2)] to the IMPACC study protocol. OHRP concurred that the study satisfied criteria for the public health surveillance exception, and the IMPACC study team sent the study protocol and participant information sheet for review and assessment to institutional review boards (IRBs) at participating institutions. Concurrent enrollment in other IRB-approved observational or interventional protocols is allowed at the discretion of the investigators and local review boards. Most participating sites elected to conduct the study under the public health surveillance exclusion; several sites with previous IRB-approved biobanking protocols elected to integrate and conduct IMPACC under those protocols.

## MATERIALS AND METHODS

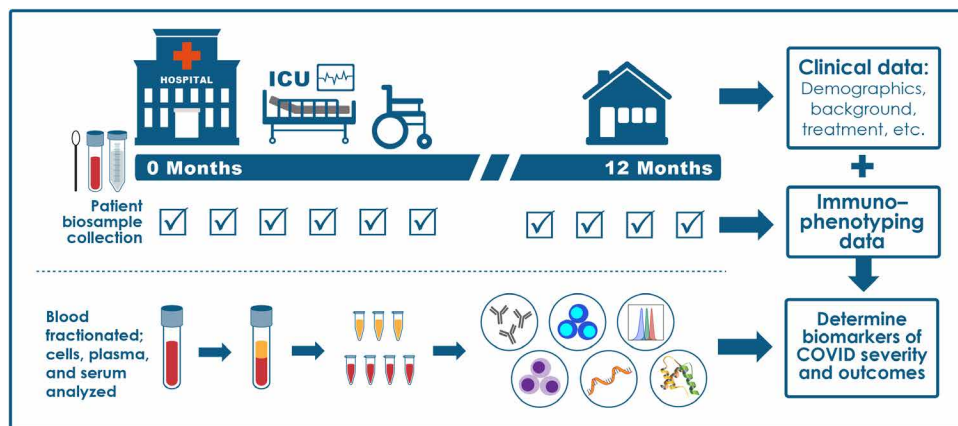
### Sample collection and processing

Sample collection was designed to meet minimal risk guidelines for blood collection for hospitalized adults, and sample-sparing assays are used when feasible. Blood samples (10 ml per time point) and nasal swabs (mid-turbinate) are collected at each specified time point, and blood is processed within 6 hours of collection according to the IMPACC standardized operating procedure. Whole blood and peripheral blood mononuclear cells (PBMCs) are collected to identify distinct immune cell populations and quantify changes in cell populations, gene expression, and activation markers [e.g., cytometry by time-of-flight (CyTOF) and bulk RNA transcriptomics] over the course of COVID-19 and convalescence.

DNA is collected from whole blood at a single time point for genetic analyses (e.g., whole-exome sequencing). Serum is used to characterize SARS-CoV-2-specific antibodies, including virus neutralization, both serum and plasma are used for proteomics and metabolomics, and plasma is used to measure soluble inflammatory mediators (e.g., cytokines and chemokines) using oligonucleotide-linked antibody detection (Olink). RNA from the nasal swab is used to assess SARS-CoV-2 viral load and genomic sequence and to evaluate changes in immune-related upper airway epithelial gene expression (i.e., bulk transcriptomics). In addition, EAs are collected from intubated patients and processed within 2 hours of collection according to the



**Fig. 1. IMPACC sites and Core Labs.** The 15 IMPACC clinical sites are located within 12 states across the United States. Core Labs are located at seven sites, six of which are co-located with clinical sites. Assays conducted by each Core Lab are indicated in Fig. 4.



**Fig. 2. IMPACC study overview.** This schematic represents the IMPACC study design in which clinical data, mid-turbinate nasal swabs, and blood samples are collected at each indicated visit during and after hospitalization. EAs are only collected from intubated patients.

**Box 1. Demographic, clinical, laboratory, and radiographic assessments.**

1. Demographics
2. Targeted medical history
3. Outpatient and inpatient medications (including experimental medications, vasopressors, and neuromuscular blockade agents)
4. COVID-19 symptoms, symptom onset, and exposure history
5. Date of admission to hospital, date of admission or transfer to ICU (if applicable), and date of discharge
6. Targeted physical findings
7. Vital signs (temperature, heart and respiration rates, oxygen saturation)
8. Chest imaging findings
9. Laboratory findings
  - a. CBC with differential
  - b. Metabolic panel (to include serum creatinine, total bilirubin, liver function tests, and electrolytes)
  - c. SpO<sub>2</sub>, arterial blood gas data
  - d. PT/INR, D-dimer
  - e. Ferritin, procalcitonin, LDH, CRP, cytokine panel
  - f. Troponin, cardiac enzymes
10. Requirement for respiratory support
  - a. New supplemental oxygen requirement (FIO<sub>2</sub> and mode of delivery)
  - b. Requirement for mechanical ventilation (include mode and settings, Pplat if available)
  - c. Requirement for ECMO
  - d. Use of prone positioning, inhaled nitric oxide
11. Requirement for new renal replacement therapy
12. Glasgow Coma Scale (GCS)
13. AVPU Scale (alert, verbal, pain, unresponsive)
14. Sequential Organ Failure Assessment Score (SOFA)

IMPACC standardized operating procedures. EA cells are assessed by CyTOF and bulk transcriptomics to identify and quantify changes in gene expression and activation state of distinct immune cell populations in the lower respiratory tract. Processed samples are barcoded and centrally tracked on a laboratory information management system (LDMS; Frontier Science). All supplies necessary for sample collection and sample processing are centrally procured and supplied to the participating sites. Sample collection, processing, and storage procedures (Fig. 3) are standardized across sites, and samples are transported to centralized core laboratories (Core Labs) in batches for testing and analysis. The complete sample processing manual of procedures is included as Supplementary Methods.

**Core Labs and technologies**

IMPACC is using a systems immunology approach for simultaneous immune profiling on the same patient sample, using a wide range of assays, during COVID-19 disease and resolution. Profiling assays, which are sample sparing and provide a comprehensive, unbiased assessment of immunologic changes in the airways and circulation during disease progression and resolution, were chosen (Fig. 3). Core immune profiling laboratories have been established for in-depth sample analysis using optimized assays. The Core Labs work in close collaboration with the IMPACC clinical sites and their respective sample processing laboratories to ensure uniformity in sample collection and processing using the standardized manual of procedures. The shared methods and reagents promote high-

quality assessment across all sample types. Each assay is performed at a single expert Core Lab or two harmonized Core Labs using rigorous standardized procedures, validated instruments and reagents, relevant controls, sample and batch randomization, assay timetables, and data sharing. The Core Lab assays have been chosen to provide comprehensive immune assessment while minimizing the amount of sample needed per assay.

**Virus detection by real-time PCR assay**

Viral shedding is increased and prolonged in severe compared with mild COVID-19 (8), but the precise relationship between viral kinetics and key clinical outcomes is not fully understood (9). Relating longitudinal kinetics of viral shedding to both host immune responses and clinical outcomes will provide important insights into mechanisms of severe disease. For patient comfort and consistency of collection, mid-turbinate swabs (rather than nasopharyngeal) are collected for serial viral quantification. Swabs are collected and placed in 1 ml of Zymo-DNA/RNA shield reagent (Zymo Research) and shipped to the Core Lab at Benaroya Research Institute (BRI). RNA is extracted using the Quick DNA-RNA MagBead Kit (Zymo Research) following the manufacturer's instructions. Reverse transcription PCR (RT-PCR) for SARS-CoV-2 is performed on RNA extracts using the SARS-CoV-2 (2019-nCoV) CDC qPCR Probe Assay with target genes 2019-nCoV\_N1, 2019-nCoV\_N2, and control human RNase P (Integrated DNA Technologies) (10, 11). Viral levels will be modeled longitudinally in relation to the kinetics of host immune responses and clinical outcomes.

**Viral sequencing**

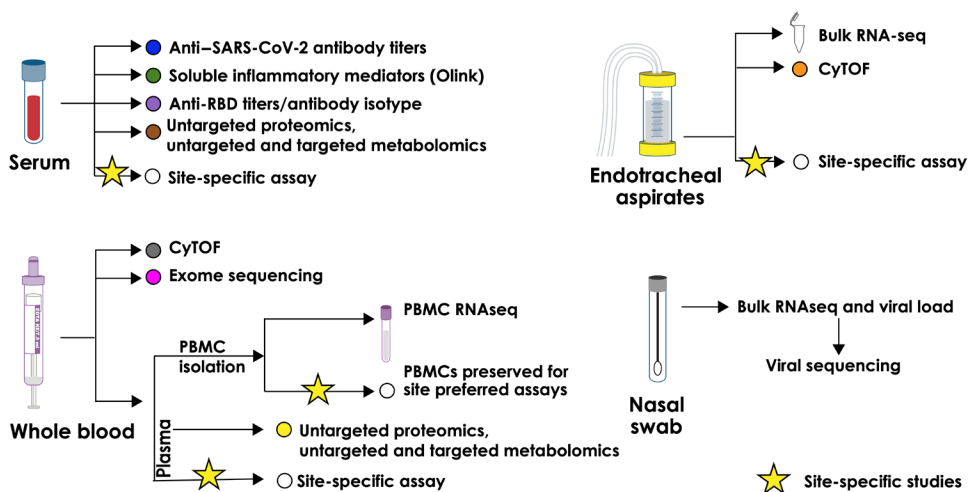
During the global spread of SARS-CoV-2, several dominant lineages have emerged that are characterized by distinct mutation patterns and are tracked in related nomenclatures by Global Initiative on Sharing All Influenza Data (GISAID) (12) and others (13). Although most genetic changes in SARS-CoV-2 are expected to be neutral, selected mutations alter viral properties such as the ability to infect cells or evade host antibody responses (14–16). Treatment with antiviral drugs can also create selective pressure for the emergence of variants that escape drug effects. The availability of complete viral genomes for all participants will allow an assessment of host responses and outcomes in the context of specific SARS-CoV-2 strains and mutations. SARS-CoV-2 viral load will be assayed directly from mid-turbinate swab samples at each time point, and viral sequencing will be performed for the first nasal swab sample per participant with sufficient quantity of virus, at the Icahn School of Medicine at Mount Sinai (ISMMS; New York, NY). The team will use a combination of tiling primer designs to specifically amplify the SARS-CoV-2 genome and generate paired-end 2 × 150-nucleotide sequencing data on the Illumina platform to an average depth of >1000-fold. If needed, supplemental data may be generated on other platforms such as Ion Torrent to maximize the ability to obtain complete genomes for patients with low viral loads. The availability of deep sequencing data from each genome will also facilitate analyses of intrahost variants.

**Serology**

Antibody responses to SARS-CoV-2 are thought to be beneficial because they can neutralize viral entry and clear infected cells through effector functions. Antibodies may also protect from reinfection, although this hypothesis is still under investigation. However, the timing and magnitude of the antibody response have been linked to disease severity (9) and certain types of antibody responses may potentially be harmful, as has been hypothesized with regard to disease-enhancing antibodies in the context of dengue infection (17).



## OVERALL PROCESSING PIPELINE



**Fig. 3. Sample processing pipeline and Core Lab assays.** Nasal swabs are used to measure viral titers, for viral genome sequencing and metagenomics, and for bulk nasal transcriptomics. Serum samples are used to measure inflammatory markers (Olink), anti-SARS and human CoV antibodies, and for untargeted proteomics and untargeted/targeted metabolomics. Whole blood is used for genome-wide association study (GWAS), WES, and CyTOF. Plasma is used for untargeted proteomics and untargeted/targeted metabolomics. PBMCs are used for bulk transcriptomics analysis. EAs are processed for bulk transcriptomics and CyTOF analyses.

The IMPACC Serology Core Lab at ISMMS is quantifying the antibody response, including different isotypes, to the Spike protein of SARS-CoV-2 and against the receptor binding domain (RBD), which is the part of the Spike protein that interacts with angiotensin-converting enzyme 2 (ACE2)-immunosorbent assays (ELISAs) (18, 19). ELISA titers will be reported as endpoint titers. Anti-spike and, especially, anti-RBD antibodies have been linked to virus neutralization. To assess the functional specificity of the antibody response, neutralization assays with authentic SARS-CoV-2 are being conducted (20). The readout for this assay is the 50% inhibitory dilution ( $ID_{50}$ ), which is calculated on the basis of virus neutralization compared with a negative control.

In addition to protein-based ELISAs, the IMPACC Serology Core Lab at UCSF is measuring serum antibody responses with a programmable phage display library [i.e., VirScan (21)] containing 38-amino acid overlapping peptides tiling across the SARS-CoV-2 and other human coronavirus (HuCoV) proteomes, including SARS-CoV-1 (NC\_004718), beta coronavirus England 1 (NC\_038294), HuCoV 229E (NC\_002645), HuCoV HKU1 (NC\_006577), HuCoV NL63 (NC\_005831), HuCoV OC43 (NC\_006213), Infectious Bronchitis virus (NC\_001451), and Middle East respiratory syndrome-related coronavirus (MERS CoV) (NC\_019843) (22). The VirScan assay measures the specificity of antibodies induced by SARS-CoV-2 infection, as well as HuCoV antibodies present early in the disease course likely as a result of past infections with other HuCoVs. These data may help determine whether the presence of preexisting HuCoV antibodies affects COVID-19 disease severity and whether specific aspects of the adaptive immune response to SARS-CoV-2 infection correlate with disease outcome. The oligonucleotides are cloned into T7 bacteriophage so that the viral peptides are displayed on the phage surface. The phage display library is then incubated with patient sera, and immunoglobulins are immunoprecipitated

by magnetic protein A and G beads along with antibody-bound phage, which are sequenced to generate viral peptide counts whose fold enrichment is calculated relative to bead-only negative controls and prepandemic healthy control sera.

### Serum cytokines

Selected proinflammatory cytokines in serum or plasma correlate with or even predict disease severity in COVID-19 (1, 23, 24). To comprehensively measure serum cytokines in a high-throughput manner for the IMPACC study, the Olink Inflammation Panel is used at the co-Core Labs at ISMMS and Stanford University. The Olink multiplex immunoassay offers the advantage of being a more comprehensive predictor of biological mechanisms at play than single cytokines (1). This method uses proximity extension assay (PEA) technology, whereby two antibodies for each target protein are conjugated to complementary oligonucleotides; if the correct antibody pair binds to the same target molecule, there is annealing of the oligos, and a PCR template can be created by

extension and then dissociation of the extended product. Quantitative PCR is then carried out on the Fluidigm Biomark microfluidic platform. This platform allows for rapid setup of 96 samples  $\times$  96 reactions, with a minimal sample requirement, nominally 1  $\mu$ l per sample. The inflammation panel consists of 92 analytes including pro- and anti-inflammatory cytokines, chemokines, and related molecules (<https://olink.com/products/inflammation>).

### Proteomics and metabolomics

The blood serves a major role in modulating and distributing the immune responses throughout the entire body. Thus, determining how soluble immunomodulatory molecules are affected by SARS-CoV-2 infection (25) and recovery from COVID-19 is essential for a comprehensive understanding of immunophenotypes. To this end, selective quantitative maps of the plasma proteome and metabolome are being acquired by the Proteomics/Metabolomics Core (PMC) at Boston Children's Hospital. To support the overarching goal of generating well-founded hypotheses to inform future research, unbiased liquid chromatography-mass spectrometry (LC-MS) methods are being used for proteomics and metabolomics. Our integrated sample-sparing proteomic and metabolomic workflow has been successfully used in populations with limited blood volumes (26).

The plasma proteomics analysis will take a two-pronged approach (27). First, the plasma proteome will be quantitatively mapped without any depletion in light of the important immunomodulatory roles of a sizable fraction of the standard depletion targets such as immunoglobulins and complement pathway components (28, 29). Mapping the COVID-19-associated changes in abundance for these proteins is important for the comprehensiveness of the immunophenotyping efforts of IMPACC. This analysis will be performed in a high-throughput fashion using a triple quadrupole mass spectrometer operated in multiple reaction mode. We target 300 of the immunologically most relevant plasma proteins

using a fast and highly sensitive state-of-the-art triple quadrupole mass spectrometer (LCMS-8060, Shimadzu). Then, the most abundant plasma samples will be depleted using biochemical methods that can be conducted in a high-throughput and cost-efficient manner on thousands of samples (30, 31). The resulting depleted plasma samples are processed for analysis by LC-MS in discovery mode using a high-throughput sample delivery and high-performance liquid chromatography (HPLC) system (Evosep One) front-end and a Bruker ion mobility/quadrupole/TOF mass spectrometer (timsTOF Pro) back-end to ensure robustness.

The plasma metabolomics for the IMPACC study also follows a two-pronged approach. First, discovery metabolomics will be conducted in collaboration with Metabolon using reversed-phase LC-MS/MS in positive ion mode, reversed-phase LC-MS/MS in negative ion mode, and hydrophilic interaction liquid chromatography (HILIC) LC-MS/MS in negative ion mode (32). In the subsequent step, select subsets of metabolites and metabolite families, deemed to be of relevance based on the discovery experiments, will be precisely quantified in a targeted fashion by HILIC LC/MS using high-accuracy/high-resolution Orbitrap mass spectrometers (Q Exactive).

#### **Transcriptional profiling (bulk RNA sequencing)**

Transcriptional profiling is a powerful approach to identify biomarkers and mechanisms of immune-mediated diseases (33, 34). Transcriptional profiling accurately reflects both dynamic changes in cellular composition and cellular response during the course of disease. Furthermore, network analysis approaches including cell deconvolution (35) and modular analysis (36) have been developed as robust computational approaches to unravel biologically coherent and insightful signatures across infectious and immunologic diseases and are particularly powerful approaches for longitudinal analyses (37–40). Transcriptional profiling of patients with COVID-19 has shown alterations in interferon responses and inflammatory pathways, which may relate to disease outcomes (35–38), sparking interest in immunomodulatory treatments directed at such pathways. The IMPACC network includes bulk transcriptomic analysis of upper (nasal, BRI Core Lab) and lower airway (EA—UCSF Core Lab) and PBMCs (Emory University and UCSF co-Core Labs).

#### **Airway bulk RNA sequencing**

RNA extracted from nasal and EA specimens is deoxyribonuclease (DNase)-treated, and human cytosolic and mitochondrial ribosomal RNA is depleted. Complementary DNA (cDNA) synthesis uses a random hexamer approach to capture human coding and noncoding RNA transcripts, as well as nonhuman RNAs. cDNA libraries are sequenced with paired-end reads at a target depth of 50 million reads per sample using NovaSeq S4 200 cycle flow cells. Human reads are aligned to the GRCh38 reference genome and quality-controlled by total counts per library and median coefficient of variation (CV) coverage. Raw counts for these genes are normalized across libraries according to the “Trimmed Means of M Values” (TMM) method (41), as implemented in the edgeR package for downstream analysis. Remaining nonhuman reads are aligned using the National Center for Biotechnology Information (NCBI) nucleotide and nonredundant protein databases followed by assembly of the reads matching each taxon detected. Repeated control and participant samples are sequenced within each sequencing batch to mitigate batch effects throughout the study.

#### **PBMC bulk RNA sequencing**

RNA is extracted from PBMCs lysed in Qiagen RLT buffer using the Zymo Quick-RNA MagBead Kit on an automated liquid handling

system in batches designed to balance covariates across library preparation runs. Library preparation is performed using the TECAN NuGen Universal Plus mRNA-seq Kit in combination with the Qiagen FastSelect hemoglobin and ribosomal depletion kit to produce stranded, poly(A)-enriched mRNA-seq libraries depleted of ribosomal and hemoglobin RNA. Libraries are normalized using the output of shallow quality control (QC) sequencing runs and pooled using appropriate dilution ratios. The targeted read depth for each sample is at least 25 million reads per sample using NovaSeq S4 200 cycle flow cells and a 100-base pair, paired-end read length. Reads are aligned to a composite reference of the GRCh38 human genome and SARS-CoV-2 genome. Gene counts are generated internally with Spliced Transcripts Alignment to a Reference (STAR), and alignments are run through RNA-Seq by Expectation Maximization (RSEM) for comparison and alternate abundance metrics; in parallel, Kallisto will be run to produce transcript per million (TPM) values (42–44). Repeated measures of healthy control and participant samples are included to track intersite batch effects; Universal Human References RNA (UHRR) controls to assess intrasite variation, and reference PBMCs stimulated with Toll-like receptor 7 (TLR7) agonist will be included to assess intrasite batch sensitivity.

#### **Mass cytometry—Blood**

Single-cell technologies have provided techniques that can resolve disease in humans at an unprecedented level of detail, capturing the clinical and biological heterogeneity of disease. Mass cytometry or CyTOF uses rare metal isotope-conjugated antibodies for high-dimensional single-cell analysis. By using heavy metal ions as labels and detection in a TOF mass spectrometer, up to 50 single-cell parameters can be measured simultaneously with little/no background and minimal signal overlap between channels, providing unprecedented multidimensional cell profiling. CyTOF has been applied in suspension to characterizing immune cells in autoimmunity, cancer, and infection (45).

The CyTOF workflow implemented for this study has been designed with several specific considerations to reduce sample usage, streamline sample processing, and minimize experimental variability (46). Whole-blood samples are first stained at the site of collection using a commercial lyophilized 30-marker panel designed to identify all major circulating immune cell subsets (Fluidigm Maxpar Direct Immune Profiling Assay, see table S1) and are then fixed and cryopreserved (Smart Tube buffer). The cryopreserved samples are shipped to the IMPACC co-Core Labs at ISMMS and Stanford for barcoded batched processing, where they are labeled with a supplemental panel of 14 additional antibodies targeting fixation-resistant epitopes to resolve additional dynamic changes in cell phenotype. To maximize reproducibility, the supplemental panel has been formulated as a cocktail and frozen in single-use aliquots for each processing batch. The labeled barcoded samples are frozen for batched acquisition. The resulting FCS files are evaluated using a centralized data processing pipeline including bead-based sample QC and data normalization and automated sample demultiplexing.

#### **Mass cytometry—EAs**

CyTOF has recently been used to analyze cells in induced sputum, and numerous studies have validated this platform for multiparameter profiling of single cells from heterogeneous populations (47). The Yale Core Lab uses CyTOF on EAs of patients with COVID-19 to provide a higher-resolution understanding of the inflammatory responses in the affected tissue.

EAs from patients who require invasive mechanical ventilation are collected at the same time points as other samples. Saline is

instilled ( $10\text{ cm}^3$ ) to collect the aspirate in a  $40\text{-cm}^3$  Argyle specimen trap. To maximize cell viability, the aspirate is processed within 2 hours of collection. Cells are passed through a series of filters to isolate a single-cell suspension for labeling by CyTOF. For optimal detection of markers, surface antigens are labeled on fresh cells with a premade batch-prepared antibody cocktail before freezing at  $-80^\circ\text{C}$ . To reduce variation, the remaining intracellular labeling is conducted on batches of samples together at the Yale University IMPACC EA CyTOF Core Lab. Antibody labeling of EA includes spiked-in reference cells (48) and markers to define polymorphonuclear leukocyte (PMN), monocytes, dendritic cells, natural killer cells, subsets of T and B lymphocytes, and 15 intracellular markers to quantify functional status (see table S2). CyTOF files from aspirates are normalized using the same centralized data processing pipeline and QC pipeline used for the CyTOF whole-blood samples, followed by a standard gating strategy for airway cells (49). EA samples are analyzed with CyTOF data from whole-blood samples of the same subjects, in concert with the other CyTOF Core teams.

### Genomic analysis

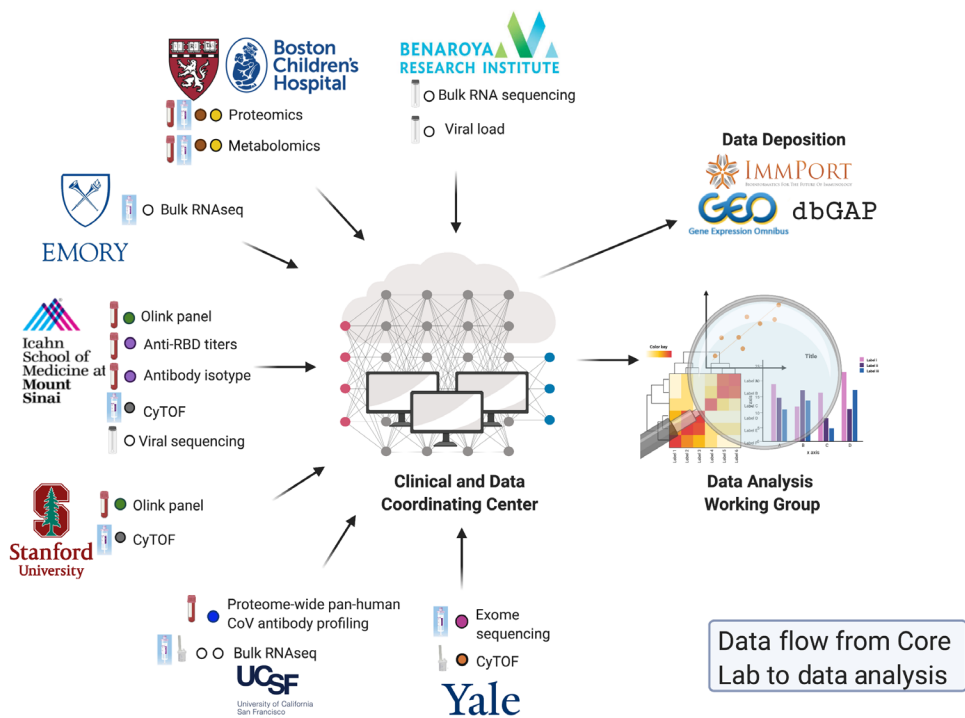
The mechanisms by which an infection leads to severe disease in a subset of all infected individuals is incompletely explained. Immune responses to infection can differ based on both rare and common genetic variations (50). To identify any genomic determinants of severe COVID-19 disease, the Yale IMPACC Core Lab is conducting whole-exome sequencing (WES) and single-nucleotide polymorphism (SNP) genotyping and assessing genetic variants associated with individual susceptibility to severe disease. The DNA sequencing of IMPACC study subjects includes WES to include 19,433 genes that are in the RefSeq coding sequences, xGen exome capture, and whole-genome genotyping at 1.9 million SNP sites on the Illumina Infinium Global Diversity Array (GDA). For WES, genomic DNA will be extracted from frozen whole blood of each enrolled subject, with sample quality determined by spectroscopic and fluorometric methods. High-quality DNA will be sheared for automated library construction incorporating unique dual indices for each sample followed by hybridization-based enrichment of the exome. Pooled libraries will be sequenced on Illumina NovaSeq6000 S4 flow cells using optimized conditions for concentrations to maximize unique read output while limiting duplicates using paired-end sequencing chemistry and a read length of 101 bases. Following real-time analysis on Illumina's CASAVA 1.8.2 software suite for converting signal intensities to individual base calls and completion of the run, raw data are evaluated for quality and samples are demultiplexed. Individual sample-level alignment to the human genome, variant calling, and annotation enable downstream analyses. Whole-genome genotyping will be performed following the manufacturer's recommendations.

Sequencing and array data will be available as FASTQ files to the analysis team accompanied by common variant association tests and rare variant gene burden tests for outcomes. Genetic sequence data will estimate population stratification and relatedness in our samples as covariates in other analyses.

## DATA MANAGEMENT AND ANALYSIS PLAN

### Study reporting and coordination

The study coordination and project management support of all participating IMPACC clinical, local laboratory processing, and Core research laboratory sites are centralized at the Clinical and Data Coordinating Center (CDCC) level (Fig. 4). A web-based clinical database system built on REDCap (51) (<https://project-redcap.org>) captures clinical data from all participants at all specified time points. CRFs harmonize with existing data standards put forward by the ISARIC, World Health Organization (WHO), and Prevention and Early Treatment of Acute Lung Injury (PETAL) Network. The clinical database contains participant- and visit-level identifiers to enable linkage with the sample tracking system and other immunophenotypic data. In coordination with the clinical information captured, a web-based sample tracking database LDMS (<https://ldms.org>) captures sample locations via barcode system in real time at point of collection. CDCC biostatistical support generates periodic reports of study enrollment and accrual, queries, summaries of primary study endpoints, final report, and any statistical tables, figures for scientific reports, and presentations. The CDCC provides training for each activity and serves as primary facilitator for intrastudy communications. This includes maintenance



**Fig. 4. Data flow from Core Lab to data analysis.** IMPACC Core Labs conduct the designated assays, QC, and preliminary data analysis before sending the validated datasets to the IMPACC CDCC for QA. After QC/QA, these data are provided to the DAWG for independent and integrated analyses to identify correlates of clinical outcomes, COVID-19 disease severity and progression, and multiomic signatures.



of a web-based communications portal to allow for individual- and group-based messaging and document sharing, a study pager to receive urgent text-based messaging, and coordination of all regularly scheduled working group meetings, which typically occur weekly, biweekly, or monthly. The CDCC provides continuous operational support, receiving and distributing any queries or concerns from study personnel related to logistics or study conduct aiming for near real-time resolution. Last, the CDCC developed and coordinates a network-wide patient newsletter to augment site strategies for participant retention.

### Collection of data and harmonization of data processing pipelines

Harmonized procedural and data processing pipelines across assays are coordinated with the Core Lab leads with the support of the Data Analysis Working Group (DAWG) and CDCC. As the cohort is recruited, plans will be made to address missing clinical data and samples and to appropriately randomize samples for each assay type. Once the data are uploaded to the IMPACC study data dashboard, CDCC data managers and biostatisticians will coordinate and verify QC processes for data collected/generated at clinical sites and Core Labs and will further perform additional quality assurance (QA) to maintain the highest possible accuracy of data before reporting and analysis.

### Centralized computing environment

The Amazon Web Services (AWS) (52) (<https://aws.amazon.com>) cloud computing platform is used for encrypted, access-controlled data storage and data analysis resources. A data and analysis dashboard allows IMPACC investigators to upload and store raw and processed computable data, and features a centralized computing environment enabling investigators to perform their analyses (53). Therefore, the AWS platform provides a computing environment for developing, testing, and running scripts, and performing QC and QA on the raw data generated by the Core Labs. This centralized computing environment also ensures that the IMPACC Core Labs use shared data standards, and are internally consistent, so that the datasets can be accurately analyzed in an integrated fashion. This approach is designed to facilitate data sharing and downstream analyses by IMPACC investigators and the broader research community, who will be able to access the data and associated metadata via the ImmPort public data repository ([www.ImmPort.org](http://www.ImmPort.org)).

### Clinical data analysis

Clinical outcomes will be defined for each patient in the study including binary outcomes (e.g., mortality yes/no) and disposition at day 14 or day 28 (e.g., discharged versus hospitalized versus deceased). Trajectory of clinical outcome will be defined at each time point that samples are collected using a seven-point ordinal scale modified from the WHO ([https://who.int/blueprint/priority-diseases/key-action/COVID-19\\_Treatment\\_Trial\\_Design\\_Master\\_Protocol\\_synopsis\\_Final\\_18022020.pdf](https://who.int/blueprint/priority-diseases/key-action/COVID-19_Treatment_Trial_Design_Master_Protocol_synopsis_Final_18022020.pdf)) and the NIAID eight-point ordinal scales (54) indicating disease severity, with clinical status defined as follows: (i) not hospitalized, no limitations; (ii) not hospitalized, activity limitations or requires home oxygen; (iii) hospitalized, not requiring supplemental oxygen; (iv) hospitalized, requires supplemental oxygen; (v) hospitalized, requires high-flow nasal cannula or noninvasive ventilation; (vi) hospitalized, requires invasive mechanical ventilation and/or extracorporeal membrane

oxygenation (ECMO); and (vii) death. Latent class analysis or other unsupervised clustering techniques will determine categories of disease severity based on clinical outcomes or trajectories. Modeling of clinical data will associate symptoms, clinical characteristics, laboratory values, radiographic findings, and other experimental assay parameters at time of hospital admission and over the course of the disease that predict these categories of disease severity.

### Planned experimental data analyses

Core Lab assay data will be analyzed to identify correlates of clinical outcomes and COVID-19 disease severity and progression. Each Core Lab assay will be analyzed independently, and later, these data will be integrated to derive multiomic signatures and mechanistic hypotheses related to patient outcomes or COVID-19 disease trajectories. The initial analysis will be carried out using established pipelines for viral loads (55–57), viral sequencing (58, 59), serology (22, 60), serum cytokines (61, 62), proteomics (27, 63), metabolomics (64, 65), transcriptional profiling (36, 39, 66), metagenomics (67, 68), mass cytometry (69, 70), and genomic (71, 72) data. In general, low-dimensional assays, such as antibody titers and viral loads, will be analyzed directly, whereas high-dimensional data, such as RNA sequencing and mass cytometry, will first be projected into a lower-dimensional space for analysis. For genomic analysis, common variant association tests and rare variant gene burden tests will be conducted for outcomes, and genetic data will be used to estimate population stratification and relatedness in the IMPACC samples as covariates in other analyses.

The IMPACC study design offers a natural means to align the longitudinal samples collected from each patient with their hospital admission date. However, this approach may not be the most informative for identifying clinical associations. The comprehensive collection of clinical data will enable assessment of alternative approaches to align the samples, including time after onset of symptoms and/or time since escalation of care. The longitudinal nature of the IMPACC study will also allow exploration of disease trajectories for individual patients. Each of the derived measurements will be examined to determine whether they are coincident, lagging, or preceding indicators of disease severity.

The availability of a wide array of Core Lab assays probing different aspects of human immunity to SARS-CoV-2 on the same samples provides the opportunity for large-scale multiomic analysis. Jointly modeling the relations among different assay measurements allows for better understanding of the interactions driving the host response to SARS-CoV-2 and how it relates to clinical outcome. A supervised approach is used first to construct sparse factor models to predict clinical outcomes (per patient) or disease severity (per sample) from multiomic measurements at each time point (73–76), as well as for the longitudinal curves (77–79). Along with supervised analyses, unsupervised multiomic approaches will be used to define groups of patients that share similar immune profiles (80–82). This clustering can be done separately for each time point or by incorporating the full longitudinal time series for each patient (36, 83–85). Once defined, the relationship between these clusters and clinical outcomes will be determined. This analysis will allow identification of both assay-specific and shared functional modules that explain the variability of clinical outcomes across patients. Identification of immunologic signatures can be validated with clinical data on disease severity and further refinement of clinical

phenotyping, resulting in an iterative process to reveal immunologic mechanisms of disease and resulting clinical manifestations. As detailed below, data will be shared to leverage innovative analysis approaches from the wider scientific community.

### Governance

The IMPACC study leadership is composed of key program officers at NIAID, key members of the CDCC and DAWG, and study site principal investigators and clinical leads. This leadership group developed several key committees and working groups to oversee the project progress, resolve conflicts, ensure effective communication within the study team to achieve the study goals, and effectively disseminate IMPACC study results to the public. The governance infrastructure includes the following committees and working groups: The Steering Committee is responsible for the overall study oversight and conflict resolution; the Core Lab Working Group reviews core assay progress, QC, data analysis, and submission of the data to the CDCC; and the Clinical Working Group in collaboration with the CDCC reviews study progress including enrollment and retention, sample collection, and clinical data. Analysis and integration of the data is overseen by the DAWG in close collaboration with the Clinical and Core Lab Working Groups. The Publications Committee ensures compliance with the publications policy and reviews and prioritizes IMPACC manuscript proposals. In addition, the CDCC meets with clinical study coordinators and processing laboratory leads at recruiting sites weekly or biweekly to identify and mitigate barriers to successful study conduct. This committee/work group structure is essential to identify any obstacles or opportunities early on and to identify and deploy resources to meet the study objectives.

### Data sharing

IMPACC data sharing follows the NIH public data sharing policy ([http://grants.nih.gov/grants/policy/data\\_sharing](http://grants.nih.gov/grants/policy/data_sharing)) to enable the widest dissemination of experimental and clinical data and associated metadata while also protecting the privacy of the participants and the utility of these datasets by deidentifying and masking potentially sensitive data elements. IMPACC will use the ImmPort database (86, 87) ([www.immport.org](http://www.immport.org)) as the primary database for experimental and clinical data and metadata dissemination. In addition, the IMPACC CDCC and DAWG leads will work with the ImmPort team to develop an IMPACC-specific resource page to facilitate public access to IMPACC datasets. Viral sequencing data will be deposited in GenBank (12, 88) to provide access to the scientific community. The CDCC is responsible for deposition of clinical and Core Lab data and associated metadata to ImmPort, whereas deposition of all published site-specific datasets will be the responsibility of individual IMPACC sites. ImmPort will work closely with CDCC and specific sites to validate, curate, and verify submitted data and metadata. The IMPACC Publication Committee provides oversight for publications by establishing a uniform policy and transparent process for the preparation, prereview, and submission of IMPACC manuscripts and datasets.

### CONCLUSION

The goals of IMPACC are to longitudinally survey a large population of hospitalized patients with COVID-19 to inform disease progression dynamics and related biomarkers and conduct detailed,

longitudinal immunophenotyping from presentation through disease progression or resolution. We aim to understand the interplay between viral load, immune pathology, and clinical manifestations of disease in a hospitalized COVID-19 cohort. Our study should improve understanding of the role of the host immune response in COVID-19 severity outcomes and help to generate hypotheses regarding effective host-directed therapeutics and optimal timing for administration of host response-directed interventions. The distinct features of IMPACC include leveraging of existing consortia and expert laboratories, efficient use of sample-sparing methods to maximize use of samples for multiple assays, and carefully harmonized data and sample collection, sample processing, and data sharing and analysis. Through this publication, we hope to facilitate coordination with other COVID-19 studies around the globe to foster data comparison and meta-analyses and to serve as a model for multicenter studies where a comprehensive, longitudinal assessment of biological processes is desired in the setting of public health emergencies.

### SUPPLEMENTARY MATERIALS

[immunology.sciencemag.org/cgi/content/full/6/62/eabf3733/DC1](https://immunology.sciencemag.org/cgi/content/full/6/62/eabf3733/DC1)

Methods: Sample Processing Manual of Procedures

Tables S1 and S2

[View/request a protocol for this paper from Bio-protocol.](#)

### REFERENCES AND NOTES

1. P. S. Arunachalam, F. Wimmers, C. K. P. Mok, R. A. P. M. Perera, M. Scott, T. Hagan, N. Sigal, Y. Feng, L. Bristow, O. Tak-Yin Tsang, D. Wagh, J. Collier, K. L. Pellegrini, D. Kazmin, G. Alaeddine, W. S. Leung, J. M. C. Chan, T. S. H. Chik, C. Y. C. Choi, C. Huerta, M. Paine McCullough, H. Lv, E. Anderson, S. Edupuganti, A. A. Upadhyay, S. E. Bosinger, H. T. Maecker, P. Khatri, N. Rouphael, M. Peiris, B. Pulendran, Systems biological assessment of immunity to mild versus severe COVID-19 infection in humans. *Science* **369**, 1210–1220 (2020).
2. A. Grifoni, D. Weiskopf, S. I. Ramirez, J. Mateus, J. M. Dan, C. R. Moderbacher, S. A. Rawlings, A. Sutherland, L. Premkumar, R. S. Jadi, D. Marrama, A. M. de Silva, A. Frazier, A. F. Carlin, J. A. Greenbaum, B. Peters, F. Krammer, D. M. Smith, S. Crotty, A. Sette, Targets of T cell responses to SARS-CoV-2 coronavirus in humans with COVID-19 disease and unexposed individuals. *Cell* **181**, 1489–1501.e15 (2020).
3. J. Hadjadj, N. Yatim, L. Barnabei, A. Corneau, J. Boussier, N. Smith, H. Péré, B. Charbit, V. Bondet, C. Chenevier-Gobeaux, P. Breillat, N. Carlier, R. Gauzit, C. Morbieu, F. Pène, N. Marin, N. Roche, T. A. Szwed, S. H. Merkle, J. M. Treliuyer, D. Veyer, L. Mouthon, C. Blanc, P. L. Tharaux, F. Rozenberg, A. Fischer, D. Duffry, F. Rieux-Laucat, S. Kernéis, B. Terrier, Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718–724 (2020).
4. D. Mathew, J. R. Giles, A. E. Baxter, D. A. Oldridge, A. R. Greenplate, J. E. Wu, C. Alano, L. Kuri-Cervantes, M. B. Pampena, M. T'okuyama, S. Manne, Z. Chen, Y. J. Huang, J. P. Reilly, A. R. Weisman, C. A. G. Ittner, O. Kuthuru, J. Dougherty, K. Nzingha, N. Han, J. Kim, A. Pattekar, E. C. Goodwin, E. M. Anderson, M. E. Weirick, S. Gouma, C. P. Arevalo, M. J. Bolton, F. Chen, S. F. Lacey, H. Ramage, S. Cherry, S. E. Hensley, S. A. Apostolidis, A. C. Huang, L. A. Vella; The UPenn COVID Processing Unit, M. R. Betts, N. J. Meyer, E. J. Wherry, Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, eabc8511 (2020).
5. C. Lucas, P. Wong, J. Klein, T. B. R. Castro, J. Silva, M. Sundaram, M. K. Ellingson, T. Mao, J. E. Oh, B. Israel, T. Takahashi, M. Tokuyama, P. Lu, A. Venkataraman, A. Park, S. Mohanty, H. Wang, A. L. Wyllie, C. B. F. Vogels, R. Earnest, S. Lapidus, I. M. Ott, A. J. Moore, M. C. Muenker, J. B. Fournier, M. Campbell, C. D. Odio, A. Casanovas-Massana; Yale IMPACT Team, A. Obaid, A. Lu-Culligan, A. Nelson, A. Brito, A. Nunez, A. Martin, A. Watkins, B. Geng, C. Kalinich, C. Harden, C. Todeasa, C. Jensen, D. Kim, D. McDonald, D. Shepard, E. Courchaine, E. B. White, E. Song, E. Silva, E. Kudo, G. Deluili, H. Rahming, H. J. Park, I. Matos, J. Nouws, J. Valdez, J. Fauver, J. Lim, K. A. Rose, K. Anastasio, K. Brower, L. Glick, L. Sharma, L. Sewanan, L. Knaggs, M. Minasyan, M. Batsu, M. Petrone, M. Kuang, M. Nakahata, M. Campbell, M. Linehan, M. H. Askense, M. Simonov, M. Smolgovsky, N. Sonner, N. Naushad, P. Vijayakumar, R. Martinello, R. Datta, R. Handoko, S. Bermejo, S. Prophet, S. Bickerton, S. Velazquez, T. Alpert, T. Rice, W. Khoury-Hanold, X. Peng, Y. Yang, Y. Cao, Y. Strong, R. Herbst, A. C. Shaw, R. Medzhitov, W. L. Schulz, N. D. Grubaugh, C. dela Cruz, S. Farhadian, A. I. Ko, S. B. Omer, A. Iwasaki, Longitudinal analyses reveal immunological misfiring in severe COVID-19. *Nature* **584**, 463–469 (2020).



6. World Health Organization, *WHO R&D Blueprint: Novel Coronavirus: COVID-19 Therapeutic Trial Synopsis* (World Health Organization, 2020).
7. E. Stolk, K. Ludwig, K. Rand, B. van Hout, J. M. Ramos-Goni, Overview, update, and lessons learned from the international EQ-5D-5L valuation work: Version 2 of the EQ-5D-5L valuation protocol. *Value Health* **22**, 23–30 (2019).
8. D. Cella, S. Yount, N. Rothrock, R. Gershon, K. Cook, B. Reeve, D. Ader, J. F. Fries, B. Bruce, M. Rose; PROMIS Cooperative Group, The Patient-Reported Outcomes Measurement Information System (PROMIS): Progress of an NIH Roadmap cooperative group during its first two years. *Med. Care* **45**, S3–S11 (2007).
9. Y. Wang, L. Zhang, L. Sang, F. Ye, S. Ruan, B. Zhong, T. Song, A. N. Alshukairi, R. Chen, Z. Zhang, M. Gan, A. Zhu, Y. Huang, L. Luo, C. K. P. Mok, M. M. al Gethamy, H. Tan, Z. Li, X. Huang, F. Li, J. Sun, Y. Zhang, L. Wen, Y. Li, Z. Chen, Z. Zhuang, J. Zhuo, C. Chen, L. Kuang, J. Wang, H. Lv, Y. Jiang, M. Li, Y. Lin, Y. Deng, L. Tang, J. Liang, J. Huang, S. Perlman, N. Zhong, J. Zhao, J. S. Malik Peiris, Y. Li, J. Zhao, Kinetics of viral load and antibody response in relation to COVID-19 severity. *J. Clin. Invest.* **130**, 5235–5244 (2020).
10. F. X. Lescure, L. Bouadma, D. Nguyen, M. Parisey, P. H. Wicky, S. Behillil, A. Gaymard, M. Bouscambert-Duchamp, F. Donati, Q. le Hingrat, V. Enouf, N. Houhou-Fidouh, M. Valette, A. Maillat, J. C. Lucet, F. Mentre, R. Duval, D. Descamps, D. Malvy, J. F. Timsit, B. Lina, S. van-der-Werf, Y. Yazdanpanah, Clinical and virological data of the first cases of COVID-19 in Europe: A case series. *Lancet Infect. Dis.* **20**, 697–706 (2020).
11. R. Wolfel, V. M. Corman, W. Guggemos, M. Seilmaier, S. Zange, M. A. Müller, D. Niemeyer, T. C. Jones, P. Vollmar, C. Rothe, M. Hoelscher, T. Bleicker, S. Brünink, J. Schneider, R. Ehmann, K. Zwirgmaier, C. Drosten, C. Wendtner, Virological assessment of hospitalized patients with COVID-2019. *Nature* **581**, 465–469 (2020).
12. Y. Shu, J. McCauley, GISAID: Global initiative on sharing all influenza data—From vision to reality. *Euro Surveill.* **22**, 30494 (2017).
13. A. Rambaut, E. C. Holmes, A. O’Toole, V. Hill, J. T. McCrone, C. Ruis, L. du Plessis, O. G. Pybus, A dynamic nomenclature proposal for SARS-CoV-2 lineages to assist genomic epidemiology. *Nat. Microbiol.* **5**, 1403–1407 (2020).
14. B. Korber, W. M. Fischer, S. Gnanakaran, H. Yoon, J. Theiler, W. Abfalterer, N. Hengartner, E. E. Giorgi, T. Bhattacharya, B. Foley, K. M. Hastie, M. D. Parker, D. G. Partridge, C. M. Evans, T. M. Freeman, T. I. de Silva; Sheffield COVID-19 Genomics Group, C. M. Danal, L. G. Perez, H. Tang, A. Moon-Walker, S. P. Whelan, C. C. La Branche, E. O. Sapphire, D. C. Montefiori, Tracking changes in SARS-CoV-2 spike: Evidence that D614G increases infectivity of the COVID-19 virus. *Cell* **182**, 812–827.e19 (2020).
15. Q. Li, J. Wu, J. Nie, L. Zhang, H. Hao, S. Liu, C. Zhao, Q. Zhang, H. Liu, L. Nie, H. Qin, M. Wang, Q. Lu, X. Li, Q. Sun, J. Liu, L. Zhang, X. Li, W. Huang, Y. Wang, The impact of mutations in SARS-CoV-2 spike on viral infectivity and antigenicity. *Cell* **182**, 1284–1294.e9 (2020).
16. A. Baum, B. O. Fulton, E. Wloga, R. Copin, K. E. Pascal, V. Russo, S. Giordano, K. Lanza, N. Negron, M. Ni, Y. Wei, G. S. Atwal, A. J. Murphy, N. Stahl, G. D. Yancopoulos, C. A. Kyrtatos, Antibody cocktail to SARS-CoV-2 spike protein prevents rapid mutational escape seen with individual antibodies. *Science* **369**, 1014–1018 (2020).
17. L. C. Katzelnick, L. Gresh, M. E. Halloran, J. C. Mercado, G. Kuan, A. Gordon, A. Balmaseda, E. Harris, Antibody-dependent enhancement of severe dengue disease in humans. *Science* **358**, 929–932 (2017).
18. F. Amanat, D. Stadlbauer, S. Strohmeier, T. H. O. Nguyen, V. Chromikova, M. McMahon, K. Jiang, G. A. Arunkumar, D. Jurczykaszak, J. Polanco, M. Bermudez-Gonzalez, G. Kleiner, T. Aydillo, L. Miorin, D. S. Fierer, L. A. Lugo, E. M. Kojic, J. Stoeber, S. T. H. Liu, C. Cunningham-Rundles, P. L. Felgner, T. Moran, A. Garcia-Sastre, D. Caplivski, A. C. Cheng, K. Kedzierska, O. Vapalahti, J. M. Hepojoki, V. Simon, F. Krammer, A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nat. Med.* **26**, 1033–1036 (2020).
19. D. Stadlbauer, F. Amanat, V. Chromikova, K. Jiang, S. Strohmeier, G. A. Arunkumar, J. Tan, D. Bhavsar, C. Capuano, E. Kirkpatrick, P. Meade, R. N. Brito, C. Teo, M. M. Mahon, V. Simon, F. Krammer, SARS-CoV-2 seroconversion in humans: A detailed protocol for a serological assay, antigen production, and test setup. *Curr. Protoc. Microbiol.* **57**, e100 (2020).
20. F. Amanat, K. M. White, L. Miorin, S. Strohmeier, M. M. Mahon, P. Meade, W.-C. Liu, R. A. Albrecht, V. Simon, L. Martinez-Sobrido, T. Moran, A. Garcia-Sastre, F. Krammer, An in vitro microneutralization assay for SARS-CoV-2 serology and drug screening. *Curr. Protoc. Microbiol.* **58**, e108 (2020).
21. G. J. Xu, T. Kula, Q. Xu, M. Z. Li, S. D. Vernon, T. Ndung’u, K. Ruxrungtham, J. Sanchez, C. Brander, R. T. Chung, K. C. O’Connor, B. Walker, H. B. Larman, S. J. Eledge, Comprehensive serological profiling of human populations using a synthetic human virome. *Science* **348**, aaa0698 (2015).
22. C. R. Zamecnik, J. V. Rajan, K. A. Yamauchi, S. A. Mann, R. P. Loudermilk, G. M. Sowa, K. C. Zorn, B. D. Alvarenga, C. Gaebler, M. Caskey, M. Stone, P. J. Norris, W. Gu, C. Y. Chiu, D. Ng, J. R. Byrnes, X. X. Zhou, J. A. Wells, D. F. Robbiani, M. C. Nussenzweig, J. L. De Risi, M. R. Wilson, ReScan, a multiplex diagnostic pipeline, pans human sera for SARS-CoV-2 antigens. *Cell Rep. Med.* **1**, 100123 (2020).
23. D. M. Del Valle, S. Kim-Schulze, H.-H. Huang, N. Beckmann, S. Nirenberg, B. Wang, Y. Lavin, T. Swartz, D. Madduri, A. Stock, T. U. Marron, H. Xie, M. Patel, O. van Oekelen, A. Rahman, P. Kovatch, J. A. Aberg, E. Schadt, S. Jagannath, M. Mazumdar, A. Charney, A. Firpo-Betancourt, D. R. Mendu, J. Jhang, D. Reich, K. Sigel, C. Cordon-Cardo, M. Feldmann, S. Parekh, M. Merad, S. Gnjatic, An inflammatory cytokine signature helps predict COVID-19 severity and death. *medRxiv* 10.1101/2020.05.28.20115758, (2020).
24. A. Mazzoni, L. Salvati, L. Maggi, M. Capone, A. Vanni, M. Spinicci, J. Mencarini, R. Caporale, B. Peruzzi, A. Antonelli, M. Trotta, L. Zammarchi, L. Ciani, L. Gori, C. Lazeri, A. Matucci, A. Vultaggio, O. Rossi, F. Almerigogna, P. Parronchi, P. Fontanari, F. Lavorini, A. Peris, G. M. Rossolini, A. Bartoloni, S. Romagnani, F. Liotta, F. Annunziato, L. Cosmi, Impaired immune cell cytotoxicity in severe COVID-19 is IL-6 dependent. *J. Clin. Invest.* **130**, 4694–4703 (2020).
25. A. D. Whetton, G. W. Preston, S. Abubeker, N. Geifman, Proteomics and informatics for understanding phases and identifying biomarkers in COVID-19 disease. *J. Proteome Res.* **19**, 4219–4232 (2020).
26. A. H. Lee, C. P. Shannon, N. Amenyogbe, T. B. Bennike, J. Diray-Arce, O. T. Idoko, E. E. Gill, R. Ben-Othman, W. S. Pomat, S. D. van Haren, K.-A. L. Cao, M. Cox, A. Darboe, R. Falsafi, D. Ferrari, D. J. Harbeson, D. He, C. Bing, S. J. Hinshaw, J. Ndure, J. Njie-Jobe, M. A. Pettengill, P. C. Richmond, R. Ford, G. Saleu, G. Masiria, J. P. Matlam, W. Kirarock, E. Roberts, M. Malek, G. Sanchez-Schmitz, A. Singh, A. Angelidou, K. K. Smolen; EPIC Consortium, R. R. Brinkman, A. Ozonoff, R. E. W. Hancock, A. H. J. van den Biggelaar, H. Steen, S. J. Tebbutt, B. Kampmann, O. Levy, T. R. Kollmann, Dynamic molecular changes during the first week of human life follow a robust developmental trajectory. *Nat. Commun.* **10**, 1092 (2019).
27. T. B. Bennike, H. Steen, High-throughput parallel proteomic sample preparation using 96-well polyvinylidene fluoride (PVDF) membranes and C18 purification plates. *Methods Mol. Biol.* **1619**, 395–402 (2017).
28. H. Keshishian, M. W. Burgess, M. A. Gillette, P. Mertins, K. R. Clauser, D. R. Mani, E. W. Kuhn, L. A. Farrell, R. E. Gerszten, S. A. Carr, Multiplexed, quantitative workflow for sensitive biomarker discovery in plasma yields novel candidates for early myocardial injury. *Mol. Cell. Proteomics* **14**, 2375–2393 (2015).
29. N. L. Anderson, N. G. Anderson, The human plasma proteome: History, character, and diagnostic prospects. *Mol. Cell. Proteomics* **1**, 845–867 (2002).
30. J. Li, L. Wang, Z. Chen, R. Xie, Y. Li, T. Hang, G. Fan, Development and validation of a rapid HPLC method for the determination of cefdinir in beagle dog plasma integrated with an automatic on-line solid-phase extraction following protein precipitation in the 96-well plate format. *J. Chromatogr. B Analyt. Technol. Biomed. Life Sci.* **895–896**, 83–88 (2012).
31. A. Zougman, J. R. Wisniewski, Beyond linker histones and high mobility group proteins: Global profiling of perchloric acid soluble proteins. *J. Proteome Res.* **5**, 925–934 (2006).
32. A. M. Evans, C. D. DeHaven, T. Barrett, M. Mitchell, E. Milgram, Integrated, nontargeted ultrahigh performance liquid chromatography/electrospray ionization tandem mass spectrometry platform for the identification and relative quantification of the small-molecule complement of biological systems. *Anal. Chem.* **81**, 6656–6667 (2009).
33. Z. Wang, M. Gerstein, M. Snyder, RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **10**, 57–63 (2009).
34. D. Chaussabel, N. Baldwin, Democratizing systems immunology with modular transcriptional repertoire analyses. *Nat. Rev. Immunol.* **14**, 271–280 (2014).
35. S. S. Shen-Orr, R. Gaujoux, Computational deconvolution: Extracting cell type-specific information from heterogeneous samples. *Curr. Opin. Immunol.* **25**, 571–578 (2013).
36. P. Langfelder, S. Horvath, WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* **9**, 559 (2008).
37. M. P. Berry, C. M. Graham, F. W. McNab, Z. Xu, S. A. A. Bloch, T. Oni, K. A. Wilkinson, R. Banichereau, J. Skinner, R. J. Wilkinson, C. Quinn, D. Blankenship, R. Dhawan, J. J. Cush, A. Mejias, O. Ramilo, O. M. Kon, V. Pascual, J. Banichereau, D. Chaussabel, A. O’Garra, An interferon-inducible neutrophil-driven blood transcriptional signature in human tuberculosis. *Nature* **466**, 973–977 (2010).
38. G. Obermoser, S. Presnell, K. Domico, H. Xu, Y. Wang, E. Anguiano, L. A. Thompson-Snipes, R. Ranganathan, B. Zeitner, A. Bjork, D. Anderson, C. Speake, E. Ruchaud, J. Skinner, L. Alsina, M. Sharma, H. Dutarte, A. Cepika, E. Israelsson, P. Nguyen, Q. A. Nguyen, A. C. Harrod, S. M. Zurawski, V. Pascual, H. Ueno, G. T. Nepom, C. Quinn, D. Blankenship, K. Palucka, J. Banichereau, D. Chaussabel, Systems scale interactive exploration reveals quantitative and qualitative differences in response to influenza and pneumococcal vaccines. *Immunity* **38**, 831–844 (2013).
39. M. C. Altman, M. A. Gill, E. Whalen, D. C. Babineau, B. Shao, A. H. Liu, B. Jepsen, R. S. Gruchalla, G. T. O’Connor, J. A. Pongratic, C. M. Kerckmar, G. K. Khurana Hershey, E. M. Zoratti, C. C. Johnson, S. J. Teach, M. Kattan, L. B. Bacharier, A. Beigelman, S. M. Sigelman, S. Presnell, J. E. Gern, P. J. Gergen, L. M. Wheatley, A. Togias, W. W. Busse, D. J. Jackson, Transcriptome network identifies mechanisms of viral and nonviral asthma exacerbations in children. *Nat. Immunol.* **20**, 637–651 (2019).

40. Y. Kotliarov, R. Sparks, A. J. Martins, M. P. Mulè, Y. Lu, M. Goswami, L. Kardava, R. Bancheureau, V. Pascual, A. Biancotto, J. Chen, P. L. Schwartzberg, N. Bansal, C. C. Liu, F. Cheung, S. Moir, J. S. Tsang, Broad immune activation underlies shared set point signatures for vaccine responsiveness in healthy individuals and disease activity in patients with lupus. *Nat. Med.* **26**, 618–629 (2020).
41. M. D. Robinson, A. Oshlack, A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* **11**, R25 (2010).
42. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016).
43. A. Dobin, C. A. Davis, F. Schlesinger, J. Drenkow, C. Zaleski, S. Jha, P. Batut, M. Chaisson, T. R. Gingeras, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
44. B. Li, C. N. Dewey, RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* **12**, 323 (2011).
45. S. Palit, C. Heuser, G. P. de Almeida, F. J. Theis, C. E. Zielinski, Meeting the challenges of high-dimensional single-cell data analysis in immunology. *Front. Immunol.* **10**, 1515 (2019).
46. D. Geanon, B. Lee, G. Kelly, D. Handler, B. Upadhyaya, J. Leech, M. Herbinet, D. D. Valle, S. Gnjatic, S. Kim-Schulze, M. Merad, A. Rahman, A streamlined CyTOF workflow to facilitate standardized multi-site immune profiling of COVID-19 patients. *medRxiv* 10.1101/2020.06.26.20141341, (2020).
47. Y. Yao, T. Welp, Q. Liu, N. Niu, X. Wang, C. J. Britto, S. Krishnaswamy, G. L. Chupp, R. R. Montgomery, Multiparameter single cell profiling of airway inflammatory cells. *Cytometry B Clin. Cytom.* **92**, 12–20 (2017).
48. K. Kleinstueber, B. Corleis, N. Rashidi, N. Nchinda, A. Lisanti, J. L. Cho, B. D. Medoff, D. Kwon, B. D. Walker, Standardization and quality control for high-dimensional mass cytometry studies of human samples. *Cytometry A* **89**, 903–913 (2016).
49. E. Stewart, X. Wang, G. L. Chupp, R. R. Montgomery, Profiling cellular heterogeneity in asthma with single cell multiparameter CyTOF. *J. Leukoc. Biol.* **108**, 1555–1564 (2020).
50. M. F. Murray, E. E. Kenny, M. D. Ritchie, D. J. Rader, A. E. Bale, M. A. Giovanni, N. S. Abul-Husn, COVID-19 outcomes and the human genome. *Genet. Med.* **22**, 1175–1177 (2020).
51. P. A. Harris, R. Taylor, R. Thielke, J. Payne, N. Gonzalez, J. G. Conde, Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J. Biomed. Inform.* **42**, 377–381 (2009).
52. V. A. Fusaro, P. Patil, E. Gafni, D. P. Wall, P. J. Tonellato, Biomedical cloud computing with Amazon Web Services. *PLOS Comput. Biol.* **7**, e1002147 (2011).
53. S. M. Vignolo, J. Diray-Arce, K. McEneaney, S. Rao, C. P. Shannon, O. T. Idoko, F. Cole, A. Darboe, F. Cessay, R. Ben-Othman, S. J. Tebbutt, B. Kampmann, P. Levy, A. Ozonoff, EPIC Consortium, A cloud-based bioinformatic analytic infrastructure and Data Management Core for the Expanded Program on Immunization Consortium. *J. Clin. Transl. Sci.* **5**, E52 (2020).
54. J. H. Beigel, K. M. Tomashek, L. E. Dodd, A. K. Mehta, B. S. Zingman, A. C. Kalil, E. Hohmann, H. Y. Chu, A. Luetkemeyer, S. Kline, D. Lopez de Castilla, R. W. Finberg, K. Dierberg, V. Tapson, L. Hsieh, T. F. Patterson, R. Paredes, D. A. Sweeney, W. R. Short, G. Touloumi, D. C. Lye, N. Ohmagari, M. D. Oh, G. M. Ruiz-Palacios, T. Benfield, G. Fätkenheuer, M. G. Kortepeter, R. L. Atmar, C. B. Creech, J. Lundgren, A. G. Babiker, S. Pett, J. D. Neaton, T. H. Burgess, T. Bonnett, M. Green, M. Makowski, A. Osinusi, S. Nayak, H. C. Lane; ACTT-1 Study Group Members, Remdesivir for the treatment of Covid-19 — Final report. *N. Engl. J. Med.* **383**, 1813–1826 (2020).
55. J. Fajnzylber, J. Regan, K. Coxen, H. Corry, C. Wong, A. Rosenthal, D. Worrall, F. Giguël, A. Piechocka-Trocha, C. Atyeo, S. Fischinger, A. Chan, K. T. Flaherty, K. Hall, M. Dougan, E. T. Ryan, E. Gillespie, R. Chishti, Y. Li, N. Jilg, D. Hanidziar, R. M. Baron, L. Baden, A. M. Tsibris, K. A. Armstrong, D. R. Kuritzkes, G. Alter, B. D. Walker, X. Yu, J. Z. Li; Massachusetts Consortium for Pathogen Readiness, SARS-CoV-2 viral load is associated with increased disease severity and mortality. *Nat. Commun.* **11**, 5493 (2020).
56. E. Pujadas, F. Chaudhry, R. McBride, F. Richter, S. Zhao, A. Wajnberg, G. Nadkarni, B. S. Glicksberg, J. Houldsworth, C. Cordon-Cardo, SARS-CoV-2 viral load predicts COVID-19 mortality. *Lancet Respir. Med.* **8**, e70 (2020).
57. M. C. Kim, C. Cui, K. R. Shin, J. Y. Bae, O. J. Kweon, M. K. Lee, S. H. Choi, S. Y. Jung, M. S. Park, J. W. Chung, Duration of culturable SARS-CoV-2 in hospitalized patients with Covid-19. *N. Engl. J. Med.* **384**, 671–673 (2021).
58. J. Klein, A. F. Brito, P. Trubin, P. Lu, P. Wong, T. Alpert, M. A. Peña-Hernández, W. Haynes, K. Kamath, F. Liu, C. B. F. Vogels, J. R. Fauver, C. Lucas, J. Oh, T. Mao, J. Silva, A. L. Wyllie, M. C. Muenker, A. Casanovas-Massana, A. J. Moore, M. E. Petrone, C. C. Kalinich; Yale IMPACT Research Team, C. D. Cruz, S. Farhadian, A. Ring, J. Shon, A. I. Ko, N. D. Grubaugh, B. Israelow, A. Iwasaki, M. M. Azar, Case study: Longitudinal immune profiling of a SARS-CoV-2 reinfection in a solid organ transplant recipient. *medRxiv* 2021.2003.2024.21253992, (2021).
59. X. Zhang, Y. Tan, Y. Ling, G. Lu, F. Liu, Z. Yi, X. Jia, M. Wu, B. Shi, S. Xu, J. Chen, W. Wang, B. Chen, L. Jiang, S. Yu, J. Lu, J. Wang, M. Xu, Z. Yuan, Q. Zhang, X. Zhang, G. Zhao, S. Wang, S. Chen, H. Lu, Viral and host factors related to the clinical outcome of COVID-19. *Nature* **583**, 437–440 (2020).
60. R. D. Schubert, I. A. Hawes, P. S. Ramachandran, A. Ramesh, E. D. Crawford, J. E. Pak, W. Wu, C. K. Cheung, B. D. O'Donovan, C. M. Tato, A. Lyden, M. Tan, R. Sit, G. A. Sowa, H. A. Sample, K. C. Zorn, D. Banerji, L. M. Khan, R. Bove, S. L. Hauser, A. A. Gelfand, B. L. Johnson-Kerner, K. Nash, K. S. Krishnamoorthy, T. Chitnis, J. Z. Ding, H. J. McMillan, C. Y. Chiu, B. Briggs, C. A. Glaser, C. Yen, V. Chu, D. A. Wadford, S. R. Dominguez, T. F. F. Ng, R. L. Marine, A. S. Lopez, W. A. Nix, A. Soldatos, M. P. Gorman, L. Benson, K. Messacar, J. L. Konopka-Anstadt, M. S. Oberste, J. L. DeRisi, M. R. Wilson, Pan-viral serology implicates enteroviruses in acute flaccid myelitis. *Nat. Med.* **25**, 1748–1752 (2019).
61. G. L. Stein-O'Brien, R. Arora, A. C. Culhane, A. V. Favorov, L. X. Garmire, C. S. Greene, L. A. Goff, Y. Li, A. Ngom, M. F. Ochs, Y. Xu, E. J. Fertig, Enter the matrix: Factorization uncovers knowledge from omics. *Trends Genet.* **34**, 790–805 (2018).
62. K. Kvelev, E. Starosvetsky, A. Ziv-Kenet, Y. Kalugny, Y. Gorelik, G. Shalev-Malul, N. Aizenbud-Reshef, T. Dubovik, M. Briller, J. Campbell, J. C. Rieckmann, N. Asbeh, D. Rimar, F. Meissner, J. Wiser, S. S. Shen-Orr, Immune-centric network of cytokines and cells in disease context identified by computational mining of PubMed. *Nat. Biotechnol.* **36**, 651–659 (2018).
63. T. B. Bennike, M. D. Bellin, Y. Xuan, A. Stensballe, F. T. Møller, G. J. Beilman, O. Levy, Z. Cruz-Monserrate, V. Andersen, J. Steen, D. L. Conwell, H. Steen, A cost-effective high-throughput plasma and serum proteomics workflow enables mapping of the molecular impact of total pancreatectomy with islet autotransplantation. *J. Proteome Res.* **17**, 1983–1992 (2018).
64. J. Diray-Arce, M. G. Conti, B. Petrova, N. Kanarek, A. Angelidou, O. Levy, Integrative metabolomics to identify molecular signatures of responses to vaccines and infections. *Metabolites* **10**, 492 (2020).
65. R. S. Kelly, D. C. Croteau-Chonka, A. Dahlin, H. Mirzakhani, A. C. Wu, E. S. Wan, M. J. McGeachie, W. Qiu, J. E. Sordillo, A. Al-Garawi, K. J. Gray, T. F. McElrath, V. J. Carey, C. B. Clish, A. A. Litonjua, S. T. Weiss, J. A. Lasky-Su, Integration of metabolomic and transcriptomic networks in pregnant women reveals biological pathways and predictive signatures associated with preeclampsia. *Metabolomics* **13**, 7 (2016).
66. X. Wang, J. Park, K. Susztak, N. R. Zhang, M. Li, Bulk tissue cell type deconvolution with multi-subject single-cell expression reference. *Nat. Commun.* **10**, 380 (2019).
67. C. Langelier, K. L. Kalantar, F. Moazed, M. R. Wilson, E. D. Crawford, T. Deiss, A. Belzer, S. Bolourchi, S. Caldera, M. Fung, A. Jauregui, K. Malcolm, A. Lyden, L. Khan, K. Vessel, J. Quan, M. Zinter, C. Y. Chiu, E. D. Chow, J. Wilson, S. Miller, M. A. Matthay, K. S. Pollard, S. Christenson, C. S. Calfee, J. L. DeRisi, Integrating host response and unbiased microbe detection for lower respiratory tract infection diagnosis in critically ill adults. *Proc. Natl. Acad. Sci. U.S.A.* **115**, E12353–E12362 (2018).
68. K. L. Kalantar, F. Moazed, S. C. Christenson, J. Wilson, T. Deiss, A. Belzer, K. Vessel, S. Caldera, A. Jauregui, S. Bolourchi, J. L. DeRisi, C. S. Calfee, C. Langelier, Metagenomic comparison of tracheal aspirate and mini-bronchial alveolar lavage for assessment of respiratory microbiota. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **316**, L578–L584 (2019).
69. X. Liu, W. Song, B. Y. Wong, T. Zhang, S. Yu, G. N. Lin, X. Ding, A comparison framework and guideline of clustering methods for mass cytometry data. *Genome Biol.* **20**, 297 (2019).
70. L. M. Weber, M. Nowicka, C. Sonesson, M. D. Robinson, diffcyt: Differential discovery in high-dimensional cytometry via high-resolution clustering. *Commun. Biol.* **2**, 183 (2019).
71. N. R. Wray, T. Lin, J. Austin, J. J. McGrath, I. B. Hickie, G. K. Murray, P. M. Visscher, From basic science to clinical application of polygenic risk scores: A primer. *JAMA Psychiat.* **78**, 101–109 (2021).
72. E. Agerbo, B. B. Trabjerg, A. D. Børglum, A. J. Schork, B. J. Vilhjálmsson, C. B. Pedersen, C. Hakulinen, C. Albiñana, D. M. Hougaard, J. Grove, J. J. McGrath, J. Bybjerg-Grauholm, O. Mors, O. Plana-Ripoll, T. Werge, N. R. Wray, P. B. Mortensen, K. L. Musliner, Risk of early-onset depression associated with polygenic liability, parental psychiatric history, and socioeconomic status. *JAMA Psychiat.* **78**, 387–397 (2021).
73. R. Tibshirani, Regression shrinkage and selection via the lasso. *J. R. Stat. Soc. B. Methodol.* **58**, 267–288 (1996).
74. A. Singh, C. P. Shannon, B. Gautier, F. Rohart, M. Vacher, S. J. Tebbutt, K. A. Lê Cao, DIABLO: An integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics* **35**, 3055–3062 (2019).
75. T. Hastie, R. Tibshirani, Generalized additive models. *Stat. Sci.* **1**, 297–310 (1986).
76. L. Breiman, Random forests. *Mach. Learn.* **45**, 5–32 (2001).
77. H. Lin, D. Zelterman, Modeling survival data: Extending the Cox model. *Dent. Tech.* **44**, 85–86 (2002).
78. R. Tibshirani, The lasso method for variable selection in the Cox model. *Stat. Med.* **16**, 385–395 (1997).
79. B. Lau, S. R. Cole, S. J. Gange, Competing risk regression models for epidemiologic data. *Am. J. Epidemiol.* **170**, 244–256 (2009).

80. A. Y. Ng, M. I. Jordan, Y. Weiss, paper presented at the Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, Vancouver, British Columbia, Canada, 2001.
81. B. Wang, A. M. Mezlini, F. Demir, M. Fiume, Z. Tu, M. Brudno, B. Haihe-Kains, A. Goldenberg, Similarity network fusion for aggregating data types on a genomic scale. *Nat. Methods* **11**, 333–337 (2014).
82. F. Murtagh, P. Contreras, Algorithms for hierarchical clustering: An overview. *WIREs Data Mining Knowledge Discov.* **2**, 86–97 (2012).
83. C. Meng, B. Kuster, A. C. Culhane, A. M. Gholami, A multivariate approach to the integration of multi-omics datasets. *BMC Bioinformatics* **15**, 162 (2014).
84. R. Argelaguet, B. Velten, D. Arno, S. Dietrich, T. Zenz, J. C. Marioni, F. Buettner, W. Huber, O. Stegle, Multi-omics factor analysis—A framework for unsupervised integration of multi-omics data sets. *Mol. Syst. Biol.* **14**, e8124 (2018).
85. L. Guan, A smoothed and probabilistic PARAFAC model with covariates. arXiv:2104.05184 [stat.ME] (12 April 2021).
86. S. Bhattacharya, S. Andorf, L. Gomes, P. Dunn, H. Schaefer, J. Pontius, P. Berger, V. Desborough, T. Smith, J. Campbell, E. Thomson, R. Monteiro, P. Guimaraes, B. Walters, J. Wisner, A. J. Butte, ImmPort: Disseminating data to the public for the future of immunology. *Immunol. Res.* **58**, 234–239 (2014).
87. S. Bhattacharya, P. Dunn, C. G. Thomas, B. Smith, H. Schaefer, J. Chen, Z. Hu, K. A. Zalocusky, R. D. Shankar, S. S. Shen-Orr, E. Thomson, J. Wisner, A. J. Butte, ImmPort, toward repurposing of open access immunological assay data for translational and clinical research. *Sci. Data* **5**, 180015 (2018).
88. J. Hadfield, C. Megill, S. M. Bell, J. Huddleston, B. Potter, C. Callender, P. Sagulenko, T. Bedford, R. A. Neher, Nextstrain: Real-time tracking of pathogen evolution. *Bioinformatics* **34**, 4121–4123 (2018).

**Acknowledgments:** We thank K. Johnson (Boston Children's Hospital) for preparing the figures and recognize the valuable contributions of all IMPACC network members and of the patient volunteers and their families. **Funding:** This study is being supported by grants R01AI104870, R01AI132774, R01AI135803, R01AI145835, U19AI057229, U19AI062629, U19AI077439, U19AI089992, U19AI090023, U19AI118608, U19AI118610, U19AI125357, U19AI128910, U19AI128913, U54AI142766, U19AI089992, and U24AI52179 from the National Institute of Allergy and Infectious Diseases (NIAID), a part of the U.S. National Institutes of Health (NIH), and P51 OD011132, S10 OD026799 from NIH. The Precision Vaccines Program is supported, in part, by the Department of Pediatrics and Chief Scientific Officer of Boston Children's Hospital.

**Author contributions:** Lead writing team: N.R., H.M., R.R.M., J.D.-A., A.O., A.D.A., E.F.R., and P.M.B.; Core clinical protocol Working Group: P.M.B., A.D.A., C.L.H., S.S., and N.R.; Core Lab leads and contributed Core Lab descriptions: M.C.A., S.E.B., W.E., F.K., C.L., O.L., H.M., R.R.M., A.R., H.S., H.v.B., and M.R.W.; Developed the IMPACC sample processing pipeline: J.D.-A., M.C.A., S.E.B., A.D.A., W.E., F.K., C.L., O.L., H.M., R.R.M., A.R., H.S., H.v.B., A.W., J.V.R., C.R.Z., M.R.W. and P.W.; Developed and contributed the Data Management and Coordination plan and section: J.D.-A., K.M., and A.O.; Developed and contributed the Data Analysis plan and sections: L.G., S.H.K., and B. Peters. **Steering Committee contributions:** Site principal investigators: M.A.A., M.C.A., D.C., M.M.D., D.J.E., E.K.H., L.I.R.E., W.E., A.F.-S., D.A.H., M.K., O.L., W.M., J.M., R.R.M., B. Pulendran, E.F.R., M.S., and R.S.; Clinical Core leads: N.A.H., C.B., L.R.B., S.C.B., C.B.C., C.S.C., C.L.H., F.K., E.M., G.A.M., K.C.N., N.R., J.S., A.C.S., and V.S.; CDCC principal investigator: A.O.; IMPACC DAWG Leads: S.H.K. and B. Peters. **Competing interests:** F.K. is an inventor on patent applications filed by Icahn School of Medicine at Mount Sinai related to SARS-CoV-2 serological assays and Newcastle disease virus–based SARS-CoV-2 vaccines. F.K. would also like to note the following, which could be perceived as a conflict of interest: He has previously published work on influenza virus vaccines with S. Gilbert (University of Oxford), has consulted for CureVac, Merck, and Pfizer (before 2020), and is currently consulting for Pfizer, Seqirus, and Avimex. F.K.'s laboratory is collaborating with Pfizer on animal models of SARS-CoV-2, is collaborating with N. Pardi at the University of Pennsylvania on mRNA vaccines against SARS-CoV-2, and worked in the past with GlaxoSmithKline on the development of influenza virus vaccines. Two of F.K.'s mentees have recently joined Moderna. C.S.C. reports receiving during the conduct of the study: grants and personal fees from Roche/Genentech, grants and personal fees from Bayer, personal fees from Quark Pharmaceuticals, personal fees from Prometic Life Sciences, personal fees from Gen1e Life Sciences, personal fees from Vasomune Therapeutics, and personal fees from Quantum

Leap Healthcare Collaborative, outside the submitted work. S.H.K. receives consulting fees from Northrop Grumman. O.L. is a named inventor on patents related to invitro modeling of human vaccine responses and vaccine adjuvants. K.C.N. is the director of the Food Allergy Research & Education (FARE) and World Allergy Organization (WAO) Center of Excellence at Stanford University; advisor at Cour Pharmaceuticals; co-founder of Before Brands, Alladapt, Latitude, and IgGenix; National Scientific Committee member at Immune Tolerance Network (ITN) and National Institutes of Health (NIH) clinical research centers; DSMB member for NHLBI; and has filed U.S. patents for basophil testing, multifood immunotherapy and prevention, monoclonal antibody from plasmablasts, and a device for diagnostics. M.R.W. receives research grant support from Roche/Genentech. M.S. has been involved in or currently holds board and/or founding roles in Organ-1, NephroSant/KITBio and has past/present consulting/advisory/ sponsored research roles with Bristol Myers Squibb, Genentech, Astellas, Novartis, Jazz Pharma, Immucor, and Natera. **Data and materials availability:** The data and metadata generated from the IMPACC network will be deposited in the ImmPort data repository ([www.immport.org](http://www.immport.org)) as the primary method of data sharing. SARS-CoV-2 genomic sequence data will be deposited in GISAID and GenBank. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit <https://creativecommons.org/licenses/by/4.0/>. This license does not apply to figures/photos/ artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

**IMPACC Manuscript Writing Team:** Nadine Rouphael<sup>18</sup>, Holden Maecker<sup>25</sup>, Ruth R. Montgomery<sup>35</sup>, Joann Diray-Arce<sup>45</sup>, Steven H. Kleinstein<sup>3</sup>, Matthew C. Altman<sup>5</sup>, Steven E. Bosinger<sup>1</sup>, Walter Eckalbar<sup>6</sup>, Leying Guan<sup>1</sup>, Catherine L. Hough<sup>8</sup>, Florian Krammer<sup>9</sup>, Charles Langelier<sup>6,10</sup>, Ofer Levy<sup>4,16</sup>, Kerry McEnaney<sup>4</sup>, Bjoern Peters<sup>11</sup>, Adeeb Rahman<sup>9</sup>, Jayant V. Rajan<sup>6</sup>, Steven Sigelman<sup>12</sup>, Hanno Steen<sup>4</sup>, Harm van Bakel<sup>9</sup>, Alyssa Ward<sup>9</sup>, Michael R. Wilson<sup>6</sup>, Prescott Woodruff<sup>6</sup>, Colin R. Zamecnik<sup>6</sup>, Alison D. Augustine<sup>12,5</sup>, Al Ozonoff<sup>45</sup>, Elaine F. Reed<sup>135</sup>, and Patrice M. Becker<sup>12,5</sup>

**IMPACC Network Steering Committee:** Nelson Agudelo Higuita<sup>14</sup>, Matthew C. Altman<sup>5</sup>, Mark A. Atkinson<sup>15</sup>, Lindsey R. Baden<sup>16</sup>, Patrice M. Becker<sup>12</sup>, Christian Bime<sup>17</sup>, Scott C. Brakenridge<sup>15</sup>, Carolyn S. Calfee<sup>6</sup>, Charles B. Cairns<sup>18</sup>, David Corry<sup>19</sup>, Mark M. Davis<sup>2</sup>, Alison D. Augustine<sup>12</sup>, Lauren I. R. Ehrlich<sup>20</sup>, Elias K. Haddad<sup>18</sup>, David J. Erle<sup>6</sup>, Ana Fernandez-Sesma<sup>9</sup>, David A. Hafler<sup>3</sup>, Catherine L. Hough<sup>8</sup>, Farrah Kheradmand<sup>19</sup>, Steven H. Kleinstein<sup>3</sup>, Monica Kraft<sup>17</sup>, Ofer Levy<sup>4,16</sup>, Grace A. McComsey<sup>21</sup>, Esther Melamed<sup>20</sup>, William Messer<sup>8</sup>, Jordan Metcalf<sup>14</sup>, Ruth R. Montgomery<sup>3</sup>, Kari C. Nadeau<sup>2</sup>, Al Ozonoff<sup>4</sup>, Bjoern Peters<sup>11</sup>, Bali Pulendran<sup>2</sup>, Elaine F. Reed<sup>13</sup>, Nadine Rouphael<sup>1</sup>, Minnie Sarwal<sup>6</sup>, Joanna Schaeffer<sup>13</sup>, Rafick Sekaly<sup>1,21</sup>, Albert C. Shaw<sup>3</sup>, Viviana Simon<sup>9</sup>

**Affiliations:** <sup>1</sup>Emory University, <sup>2</sup>Stanford University, <sup>3</sup>Yale School of Medicine, <sup>4</sup>Boston Children's Hospital-Harvard Medical School, <sup>5</sup>Benaroya Research Institute, <sup>6</sup>University of California-San Francisco School of Medicine, <sup>7</sup>Yale School of Public Health, <sup>8</sup>Oregon and Health Sciences University, <sup>9</sup>Icahn School of Medicine at Mount Sinai, <sup>10</sup>Chan Zuckerberg Biohub, <sup>11</sup>La Jolla Institute for Immunology, <sup>12</sup>National Institutes of Allergy and Infectious Diseases/ National Institutes of Health, <sup>13</sup>University of California- Los Angeles, <sup>14</sup>Oklahoma University and Health Science Center, <sup>15</sup>University of Florida, <sup>16</sup>Brigham and Women's Hospital, <sup>17</sup>University of Arizona, <sup>18</sup>Drexel University, <sup>19</sup>Baylor College of Medicine, <sup>20</sup>The University of Texas at Austin, <sup>21</sup>Case Western Reserve University

\*Corresponding authors: [augustine@niaid.nih.gov](mailto:augustine@niaid.nih.gov); [patrice.becker@nih.gov](mailto:patrice.becker@nih.gov)  
 †Lead writing team; contributed equally to the manuscript

Submitted 22 October 2020

Accepted 5 August 2021

Published First Release 10 August 2021

Final published 7 October 2021

10.1126/sciimmunol.abf3733

**Citation:** IMPACC Manuscript Writing Team on behalf of the IMPACC Network Steering Committee, Immunophenotyping assessment in a COVID-19 cohort (IMPACC): A prospective longitudinal study. *Sci. Immunol.* **6**, eabf3733 (2021).