Original Investigation | Health Informatics

International Analysis of Electronic Health Records of Children and Youth Hospitalized With COVID-19 Infection in 6 Countries

Florence T. Bourgeois, MD, MPH; Alba Gutiérrez-Sacristán, PhD; Mark S. Keller, BS; Molei Liu, PhD; Chuan Hong, PhD; Clara-Lea Bonzel, MSc; Amelia L. M. Tan, PhD; Bruce J. Aronow, PhD; Martin Boeker, MD, MCSc, MME; John Booth, MSc; Jaime Cruz Rojo, MD; Batsal Devkota, PhD; Noelia García Barrio, MS; Nils Gehlenborg, PhD; Alon Geva, MD, MPH; David A. Hanauer, MD, MS; Meghan R. Hutch, BS; Richard W. Issitt, DClinP; Jeffrey G. Klann, PhD; Yuan Luo, PhD; Kenneth D. Mandl, MD, MPH; Chengsheng Mao, PhD; Bertrand Moal, PhD; Karyn L. Moshal, MBChB, MRCPCH, MRCP; Shawn N. Murphy, MD, PhD; Antoine Neuraz, MD, PhD; Kee Yuan Ngiam, MBBS, MRCS, MMed; Gilbert S Omenn, MD, PhD; Lav P. Patel, MS; Miguel Pedrera Jiménez, MS; Neil J. Sebire, MD; Pablo Serrano Balazote, MD, MS; Arnaud Serret-Larmande, MD; Andrew M. South, MD, MS; Anastasia Spiridou, PhD, MSc; Deanne M. Taylor, PhD; Patric Tippmann, MSc; Shyam Visweswaran, MD, PhD; Griffin M. Weber, MD, PhD; Isaac S. Kohane, MD, PhD; Tianxi Cai, PhD; Paul Avillach, MD, PhD; for the Consortium for Clinical Characterization of COVID-19 by EHR (4CE)

Abstract

IMPORTANCE Additional sources of pediatric epidemiological and clinical data are needed to efficiently study COVID-19 in children and youth and inform infection prevention and clinical treatment of pediatric patients.

OBJECTIVE To describe international hospitalization trends and key epidemiological and clinical features of children and youth with COVID-19.

DESIGN, SETTING, AND PARTICIPANTS This retrospective cohort study included pediatric patients hospitalized between February 2 and October 10, 2020. Patient-level electronic health record (EHR) data were collected across 27 hospitals in France, Germany, Spain, Singapore, the UK, and the US. Patients younger than 21 years who tested positive for COVID-19 and were hospitalized at an institution participating in the Consortium for Clinical Characterization of COVID-19 by EHR were included in the study.

MAIN OUTCOMES AND MEASURES Patient characteristics, clinical features, and medication use.

RESULTS There were 347 males (52%; 95% CI, 48.5-55.3) and 324 females (48%; 95% CI, 44.4-51.3) in this study's cohort. There was a bimodal age distribution, with the greatest proportion of patients in the 0- to 2-year (199 patients [30%]) and 12- to 17-year (170 patients [25%]) age range. Trends in hospitalizations for 671 children and youth found discrete surges with variable timing across 6 countries. Data from this cohort mirrored national-level pediatric hospitalization trends for most countries with available data, with peaks in hospitalizations during the initial spring surge occurring within 23 days in the national-level and 4CE data. A total of 27 364 laboratory values for 16 laboratory tests were analyzed, with mean values indicating elevations in markers of inflammation (C-reactive protein, 83 mg/L; 95% CI, 53-112 mg/L; ferritin, 417 ng/mL; 95% CI, 228-607 ng/mL; and procalcitonin, 1.45 ng/mL; 95% CI, 0.13-2.77 ng/mL). Abnormalities in coagulation were also evident (D-dimer, 0.78 ug/mL; 95% CI, 0.35-1.21 ug/mL; and fibrinogen, 477 mg/dL; 95% CI, 385-569 mg/dL). Cardiac troponin, when checked (n = 59), was elevated (0.032 ng/mL; 95% CI, 0.000-0.080 ng/mL). Common complications included cardiac arrhythmias (15.0%; 95% CI, 8.1%-21.7%), viral pneumonia (13.3%; 95% CI, 6.5%-20.1%), and respiratory failure (10.5%; 95% CI, 5.8%-15.3%). Few children were treated with COVID-19–directed medications.

CONCLUSIONS AND RELEVANCE This study of EHRs of children and youth hospitalized for COVID-19 in 6 countries demonstrated variability in hospitalization trends across countries and

(continued)

Key Points

Question What are international trends in hospitalizations for children and youth with SARS-CoV-2, and what are the epidemiological and clinical features of these patients?

Findings This cohort study of 671 children and youth found discrete surges in hospitalizations with variable trends and timing across countries. Common complications included cardiac arrhythmias and viral pneumonia, and laboratory findings included elevations in markers of inflammation and abnormalities of coagulation; few children and youth were treated with medications directed specifically at SARS-CoV-2.

Meaning These findings suggest large-scale informatics-based approaches used to incorporate electronic health record data across health care systems can provide an efficient source of information to monitor disease activity and define epidemiological and clinical features of pediatric patients hospitalized with SARS-CoV-2 infections.

+ **Invited Commentary**

+ **Supplemental content**

Author affiliations and article information are listed at the end of this article.

Abstract (continued)

identified common complications and laboratory abnormalities in children and youth with COVID-19 infection. Large-scale informatics-based approaches to integrate and analyze data across health care systems complement methods of disease surveillance and advance understanding of epidemiological and clinical features associated with COVID-19 in children and youth.

Introduction

The clinical presentation of coronavirus disease 2019 (COVID-19) differs substantially between children and youth and adults. The unique clinical features, complications, and outcomes of COVID-19 among children and youth warrant special consideration in epidemiologic, management, and prevention studies.[1] However, the low prevalence of disease in children and youth—compounded by the routine challenges of conducting large clinical trials in pediatric populations—has limited their inclusion in many studies.[2] Key questions remain related to risk factors for severe and rare disease manifestations and optimal use of clinical interventions.[3] The experience with COVID-19 has highlighted the critical need to have efficient methods to complement traditional clinical investigations and public health surveillance to study pediatric populations during a rapidly evolving pandemic.

Large volumes of clinical data are available in electronic health records (EHRs) to support epidemiological studies of medical conditions and analyze real-world outcomes related to specific populations and interventions.[4] When used appropriately, these data represent a powerful tool to fill in gaps and address shortcomings of conventional clinical trials. For example, EHR data have been applied to more efficiently assess medication safety in children and youth or to test at scale potential associations between risk factors and pediatric conditions.[5,6] These data are particularly conducive to the study of small populations or rare events that can be difficult to capture in smaller data sets.[7,8] Other key benefits of EHR data include the ability to ascertain clinical trajectories and to facilitate multinational studies by combining data across health care systems. Soon, EHR-based observational data may also contribute to assessing the impact of vaccines in children and youth, including efficacy and long-term safety in pediatric subpopulations with limited representation or follow-up in clinical trials.

The Consortium for Clinical Characterization of COVID-19 by EHR (4CE) is an international collaborative covering 351 adult and pediatric hospitals in 7 countries that has collected patient-level EHR data on 39 200 hospitalized patients with polymerase chain reaction (PCR)–confirmed diagnosis of SARS-CoV-2.[9] The use of common data elements across a federated network allows for integration and harmonization of data to enable analyses of the disease manifestation and epidemiology of COVID-19 across health care sites. Focusing on adult populations, studies have used data from the 4CE initiative to measure the prevalence of specific types of clinical complications, develop EHR-based severity algorithms,[10] identify laboratory tests predicting severity in patients with COVID-19,[11] and define country-level differences in demographic and epidemiological presentation.[9] Leveraging data from this collaborative, our objective was to demonstrate large-scale, multinational use of EHR data to study COVID-19 in children and youth and describe hospitalization trends and key epidemiological and clinical features of the disease.

Methods

In this cohort study, each participating site obtained institutional review board approval to share deidentified, aggregated patient data with the 4CE consortium. Informed consent was waived

because the patient data were deidentified. The study followed the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) reporting guideline.

Participating Sites and Cohort Identification

Participating 4CE sites in France, Germany, Spain, Singapore, the UK, and the US reported pediatric-specific data and contributed patients to this cohort analysis. We analyzed patients younger than 21 years who were hospitalized between February 2 to October 10, 2020, and had a positive reverse transcription PCR test for SARS-CoV-2 infection 7 days before to 14 days after the date of admission. Positive tests were identified by local data managers at each site who mapped internal codes for SARS-CoV-2 laboratory results. Demographic information on a subset of patients admitted through April 11, 2020, was previously described.[9]

Data Extraction and Aggregation

Several sites included multiple hospitals, and pediatric data were extracted from each hospital participating in the pediatric substudy (eTable in Supplement 1). Certain sites applied obfuscation thresholds to minimize disclosure risks related to small patient numbers. When values were obfuscated, we inserted a value of 0.5 times the obfuscation threshold.

Sites executed queries on local clinical data warehouses containing patient-level EHR data.[9] To construct the required data files, sites used the Informatics for Integrating Biology and the Bedside (i2b2) platform, the Observational Medical Outcomes Partnership (OMOP) Common Data Model, Epic Clarity, or other clinical data warehouses. Data files consisted of 6 tables containing aggregate patient counts for demographic characteristics, clinical course, daily counts, medication class, diagnosis, and laboratory values with mean (SD) (eFigure 1 in Supplement 1). All sites reported diagnosis as *International Statistical Classification of Diseases and Related Health Problems, Tenth Revision* (*ICD-10*) codes and used logical identifier names and codes (LOINC) for laboratory tests and anatomical therapeutic chemical National Drug Codes for medications. Each contributing site uploaded their files to a central 4CE data upload tool, where quality control and validation steps were performed before analysis.[11] Patient-level files remained at each site and were not centrally shared at any point.

Demographic and Clinical Variables

Race and ethnicity data were collected by participating hospitals based on routine practices using local race and ethnicity classifications. Sites mapped these categories to the standard categories provided by the US National Institutes of Health before the file upload to 4CE.[12] We chose to assess race and ethnicity in this study because prior reports have indicated an association between race and ethnicity and clinical outcomes for children and youth with COVID-19.[13-15]

A set of 16 laboratory values were selected, reflecting laboratory tests commonly performed as well as tests reported in prior studies to be abnormal in patients with COVID-19.[16] To describe clinical complications, we analyzed all diagnostic codes assigned to patients during the hospitalization. The diagnosis codes were reported from all sites using *ICD-10*. These codes were truncated to the first 3 characters, which represent the disease category. The codes that follow the first 3 characters add more detailed information about etiology, anatomic site, or manifestations, but would have resulted in too many categories with very low counts. To assess medication use, we determined the number of patients treated with a prespecified set of medications. These included repurposed agents used to manage COVID-19 during the study period (eg, hydroxychloroquine), investigational agents (eg, remdesivir), and adjunctive therapies used to manage complications related to COVID-19.[17]

Quality-Control Procedures

Validation steps were performed to check that file and column names were correct, columns were in the correct order, values used the correct codes or were within allowed ranges, and that there were no duplicate records. An R script was run to perform additional quality control, including ensuring the

3-digit diagnosis codes were consistent with the *ICD* dictionary. Because all laboratory tests were mapped to the same LOINC codes with unified units, laboratory test values from each site were manually reviewed to ensure the result ranges were generally consistent with data observed across other sites. Sites with implausible laboratory values or values consistently lower or higher than other sites were contacted for further investigation and correction as needed.[11] The local investigations and final assessment of accepted values considered age-specific reference ranges as well as clinical assays and site ranges.

Statistical Analysis

We summarized the daily hospitalized case counts over time and the breakdown of the cases by demographic subgroups based on pooled analysis across participating hospitals by country. To describe the clinical profile of hospitalized cases, we reported mean laboratory values at admission and percentages of frequently observed complications. Mean values and percentages with 95% CIs were aggregated across all sites based on random-effects meta-analysis.[18] To summarize temporal trends of laboratory values, we combined data from sites with at least 3 observations and calculated mean laboratory values on each day of hospitalization, also using random-effects meta-analysis.[18] Additional details on this approach are provided in the eMethods and eFigures 2 and 3 in Supplement 1. We based 95% CIs on the *z*-statistic with normal approximations for both continuous outcomes and the proportion of binary outcomes. Statistical significance was prespecified at *P* < .05 and tests were 2-tailed.

Statistical analyses and visualizations were performed in R version 3.5.1 (R Project for Statistical Computing) and Python version 3.7 (Python). We used the Altair package[19] to create figures for static publication and interactive web-based exploration. The Structured Query Language code used for data extraction, R Code used for analysis, and mapping tables used for laboratory tests and medications are available on GitHub.

Results

Study Cohort

There were 347 male patients (52%; 95% CI, 48.5%-55.3%) and 324 female patients (48%; 95% CI, 44.4%-51.3%) in our cohort. There was a bimodal age distribution, with the greatest proportion of patients in the 0- to 2-year (199 patients [30%]) and 12- to 17-year (170 patients [25%]) age range (**Figure 1**). Race and ethnicity data were not collected by sites in France, Germany, and Spain in accordance with national practices and standards. For Singapore, only 1 option was provided for Asian race, thus all local Asian groups were classified as Asian patients.

Data were collected on 671 hospitalized children and youth with PCR-confirmed SARS-CoV-2 infection across a total of 27 hospitals in 6 countries, including in France (4 hospitals), Germany (1 hospital), Singapore (1 hospital), Spain (1 hospital), the UK (1 hospital), and the US (19 hospitals). Pediatric cases were identified at each site, with France contributing 145 cases, Germany, 8 cases; Singapore, 24 cases; Spain, 78 cases; the UK, 62 cases; and the US, 354 cases.

Figure 2 illustrates the number of hospitalized pediatric patients by date during the study period.[20-24] Trends demonstrated discrete surges in hospitalization counts, with most countries experiencing a distinct increase in pediatric hospitalizations in the spring followed by variations in the occurrence and timing of subsequent surges. National-level data on pediatric hospitalizations for France, Germany, Spain, and the UK mirrored the 4CE hospitalization data. For example, peaks during the spring of 2020 in national-level data and 4CE hospitalizations occurred 19 days apart in France and 11 days apart in Germany. The largest difference was in Spain, with peaks during the initial spring surge occurring 23 days apart. National-level data were not available for Singapore and only available for a 2-month period for the US. Additional visualizations, including interactive figures with cumulative counts by country are available online.[25]

Figure 1. Age and Race/Ethnicity Distribution by Country
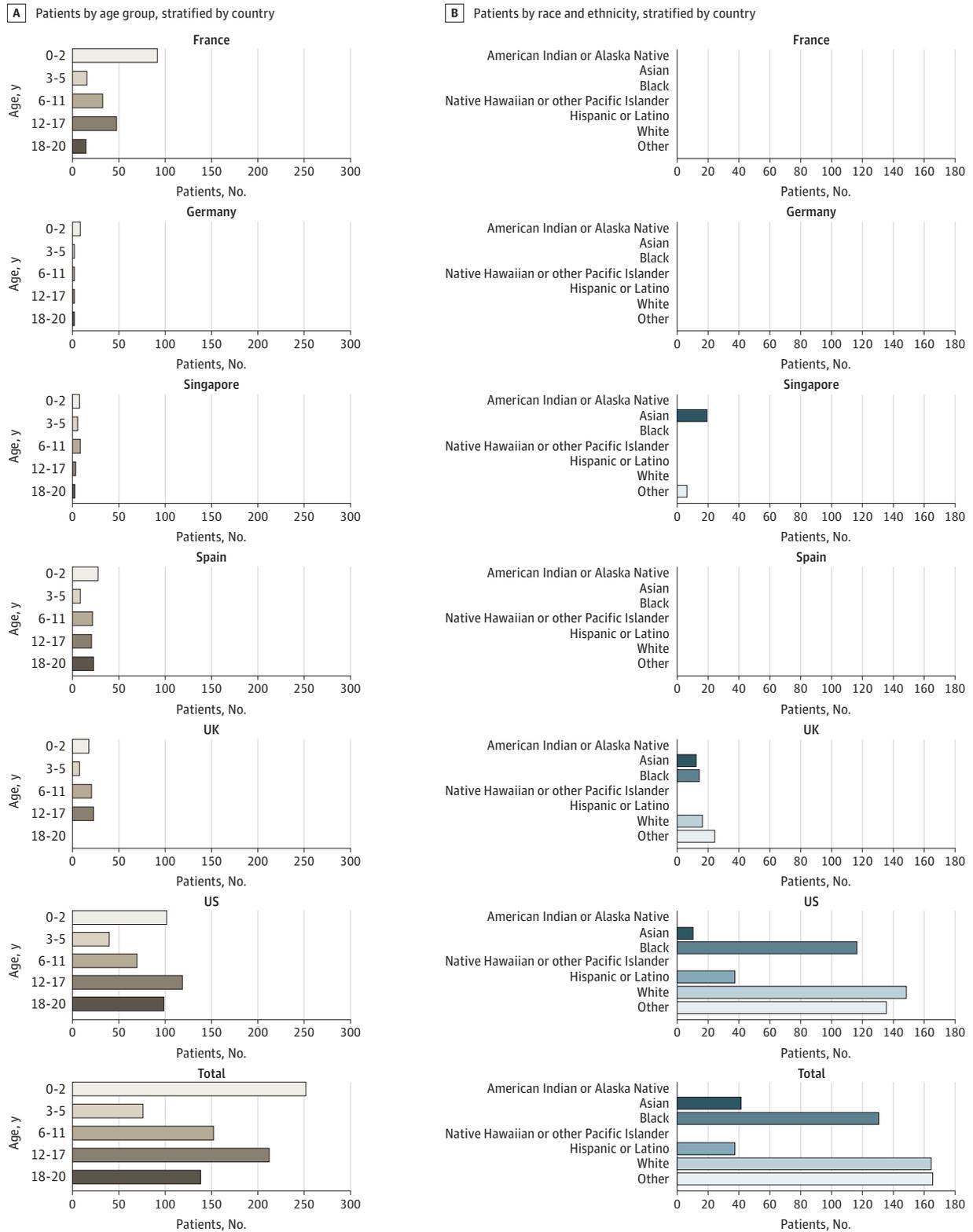
The race/ethnicity variable was based on the categories as defined by the US National Institutes of Health.[12] For Singapore, the term Asian includes Chinese, Asian Indian, and Malaysian and the term other was used for Eurasian and o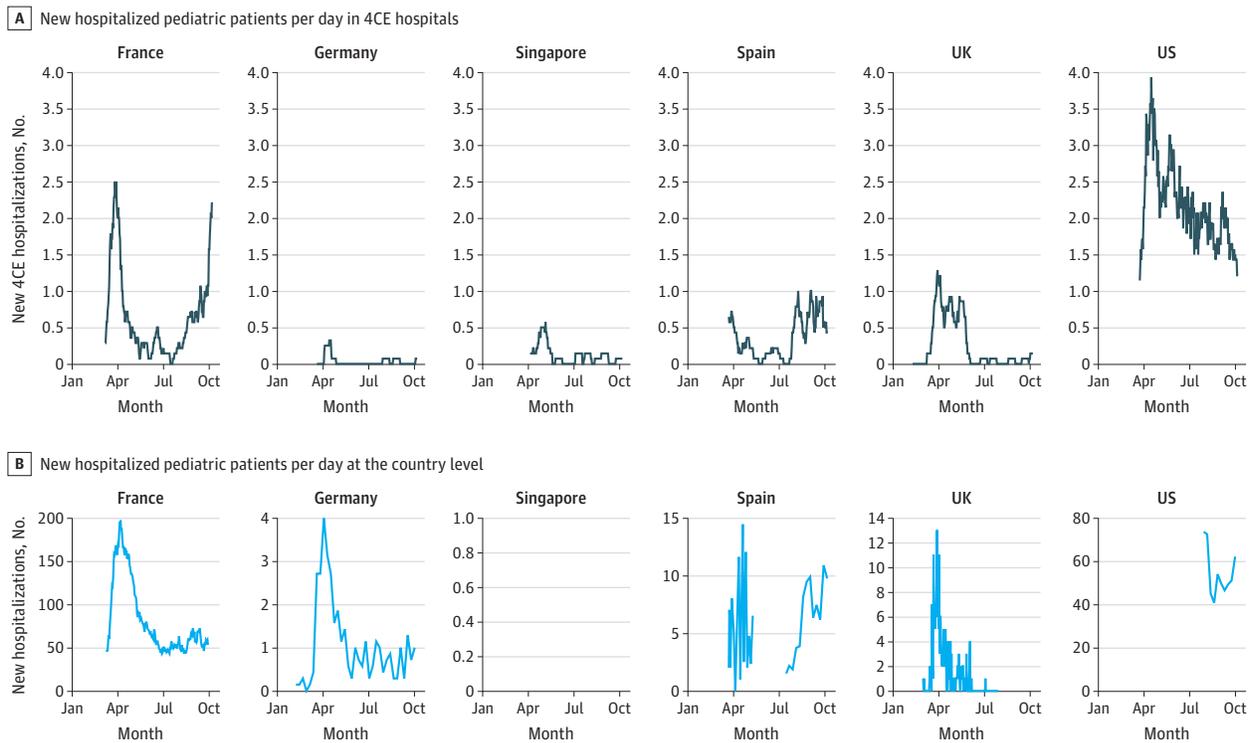ther races and ethnicities. For patients in the UK and the US, the term other represents other races and ethnicities, mixed races, and missing information on race. Information on race and ethnicity were not collected in France, Germany, and Spain.

Clinical Features

A total of 27 364 laboratory values were obtained for the 16 laboratory tests examined (**Table 1**). Mean values across hospitals were abnormal at the time of admission for markers of inflammation and coagulation. Specifically, C-reactive protein was elevated to 83 mg/L (95% CI, 53-112 mg/L; to convert to milligrams per liter, multiply by 10), ferritin to 417 ng/mL (95% CI, 228-607 ng/mL; to convert to micrograms per liter, multiply by 1), and procalcitonin to 1.45 ng/mL (95% CI, 0.13-2.77 ng/mL). However, mean values for both white blood cell count and neutrophil count were within normal limits. Dimerized plasmin fragment D (D-dimer) was elevated to 0.78 µg/mL (95% CI, 0.35-1.21 µg/mL; to convert to nanomoles per liter, multiply by 5.476) and fibrinogen to 477 mg/dL (95% CI, 385-569 mg/dL; to convert to grams per liter, multiply by .01). In a subset of patients (n = 59), cardiac troponin was elevated to 0.032 ng/mL (95% CI, 0.000-0.080 ng/mL; to convert to milligrams per liter, multiply by 1.0). Common complications included cardiac arrhythmias (15.0%; 95% CI, 8.1%-21.7%), viral pneumonia (13.3%; 95% CI, 6.5%-20.1%), and respiratory failure (10.5%; 95% CI, 5.8%-15.3%). The total number of deaths across participating sites was 18 (2.7%).

Figure 3 illustrates trajectories of laboratory values for the first 14 sequential days since admission. The number of patients with laboratory tests was highest during the initial days of hospitalization with subsequent decreases for all laboratory tests over the course of hospitalization (eFigure 3 in Supplement 1). Although the selective ordering of laboratory tests based on patients' condition limits interpretation, markers of inflammation (C-reactive protein, ferritin, neutrophil count, procalcitonin) that were initially elevated generally showed improvement after hospital days

Figure 2. Hospitalized Pediatric Case Counts by Country

A | New hospitalized pediatric patients per day in 4CE hospitals

B | New hospitalized pediatric patients per day at the country level

For France, daily pediatric hospitalization data were obtained from Santé Publique France.[20] For Germany, weekly pediatric hospitalization data were obtained from the German Society for Pediatric Infectious Diseases.[21] National pediatric hospitalization data were not available for Singapore. For Spain, weekly pediatric hospitalization data were obtained from the Spanish National Epidemiological Surveillance Network, which lacks hospitalization counts between May 11 and July 15, 2020.[22] For the UK, daily pediatric hospitalization data were obtained from the Royal College of Paediatrics and Child Health and represent pediatric hospitalizations in England.[23] For the US, weekly pediatric hospitalization data between July 31, 2020, and October 9, 2020, were obtained from the Department of Health and Human Services.[24] The y-axis scales for country-level data are independent to compare country-level trends with Consortium for Clinical Characterization of COVID-19 by EHR (4CE) trends. The plots in Figure 2A display the counts with a 14-day (centered) rolling mean.

2 to 4. For example, compared with the initial values for C-reactive protein, 4-day measurements showed a decrease of 18 mg/L (95% CI, −16-54 mg/L). Interestingly, there was a peak in several laboratory values, such as albumin, D-dimer, and lactate dehydrogenase, starting on hospital days 6 to 8. For example, compared with the initial values for D-dimer, 8-day measurements showed an increase of 1.45 µg/mL (95% CI, 0.59-2.31 µg/mL).

Medication Use

To examine the use of specific drug classes in the treatment of COVID-19 in children and youth, we determined the number of sites treating at least 3 patients with a range of drugs considered candidate therapeutic agents in adults during the study period or used to manage certain complications and underlying conditions potentially exacerbated by COVID-19 (**Table 2**). Only 2 sites treated at least 3 patients with an aminoquinoline, which includes hydroxychloroquine, and 1 site administered remdesivir to at least 3 patients. More sites administered adjunctive therapies, such as antithrombotic agents (8 sites), diuretics (8 sites), interleukin inhibitors (3 sites), and angiotensin converting enzyme inhibitors (3 sites).

Discussion

Using patient-level EHR data extracted from health care systems across 6 countries, this study offers insights on international trends of hospitalizations for children and youth with COVID-19 and defines epidemiological and clinical features associated with the disease in children and youth. Even among countries with few participating sites, hospitalization counts for children and youth over an 8 month period approximated population-level infection rates, demonstrating the potential application of this approach to monitoring disease activity in pediatric populations. Consistent with prior reports,

Table 1. Clinical Features of Children and Youth Hospitalized With COVID-19

| Characteristic	Values at admission, mean (95% CI)
Laboratory results[a]	
Alanine aminotransferase (reference range, 0-41 U/L)	33.2 (28.8-37.6) U/L
Albumin (reference range, 2.5-5 g/dL)	3.7 (3.4-4.0) g/dL
Aspartate aminotransferase (reference range, 8-92 U/L)	44 (34-55) U/L
Cardiac troponin (reference range, 0-0.01 ng/mL)	0.032 (0.000-0.080) ng/mL
C-reactive protein (reference range, 0-5 mg/L)	83 (53-112) mg/L
Creatinine (reference range, 0.1-1.0 mg/dL)	0.65 (0.59-0.70) mg/dL
D-dimer (reference range, 0-500 µg/mL)	0.78 (0.35-1.21) µg/mL
Ferritin (reference range, 12-200 ng/mL)	417 (228-607) ng/mL
Fibrinogen (reference range, 150-400 mg/dL)	477 (385-569) mg/dL
Lactate dehydrogenase (reference range, 400-1300 U/L)	449 (385-569) U/L
Lymphocyte count (reference range, 1.5-10 K/µL)	1.9 (1.7-2.2) K/µL
Neutrophil count (reference range, 1.5-10 K/µL)	6.2 (5.6-6.7) K/µL
Procalcitonin (reference range, 0-0.15 ng/mL)	1.45 (0.13-2.77) ng/mL
Prothrombin time (reference range, 8.5-17 s)	14.4 (13.3-15.4) s
Total bilirubin (reference range, 0-1.2 mg/dL)	0.76 (0.55-0.97) mg/dL
White blood cell count (reference range, 5-26 K/µL)	9.4 (8.7-10.0) K/µL
Complications, proportion of patients during hospitalization (95% CI)	
Cardiac arrhythmias	15.0 (8.1-21.7)
Viral Pneumonia	13.3 (6.5-20.1)
Respiratory failure	10.5 (5.8-15.3)
Shock	9.5 (3.7-15.3)
Acute kidney failure	7.2 (3.5-10.8)
Pleural effusion	6.6 (3.0-10.2)
Length of hospitalization, mean (95% CI), d	11.1 (8.1-14.2)

Abbreviation: D-dimer, dimerized plasmin fragment D.

SI conversion factors: To convert alanine aminotransferase to microkatal per liter, multiply by 0.0167; albumin to g/L, multiply by 10; aspartate aminotransferase to microkatal per liter, multiply by 0.0167; C-reactive protein to milligrams per liter; creatinine to micromoles per liter, multiply by 76.25; ferritin to micrograms per liter, multiply by 1; D-dimer to nanomoles per liter, multiply by 5.476; fibrinogen to grams per liter, multiply by .01; lactate dehydrogenase to microkatal per liter, multiply by 0.0167; lymphocyte count to proportion of 1.0, multiply by 0.01; neutrophil to proportion of 1.0, multiply by 0.01; total bilirubin to micromoles per liter, multiply by 17.104; troponin to milligrams per liter, multiply by 1.0, white blood cell count to proportion of 1.0, multiply by 0.01.

[a] Reference ranges as reported in The Royal College of Paediatrics and Child Health.

we found greater proportions of younger children among hospitalized patients.[26,27] Laboratory tests obtained on hospital admission indicated abnormalities in inflammation and coagulation. Examination of management patterns revealed that the use of candidate therapeutic agents adopted in adult populations remained low in children and youth.

Our study demonstrates the value of using routinely collected data from EHRs to complement other forms of disease surveillance, especially when disease prevalence is low and rapid progression precludes the development of prospective research infrastructures.[28] These data may be particularly valuable in advancing our understanding of COVID-19 in children and youth, where fewer resources have focused on COVID-19–related illness because of the less severe impact of the disease and much lower disease prevalence. While there are important limitations to EHR data, including inconsistent

Figure 3. Trajectories for Laboratory Values During Hospitalization
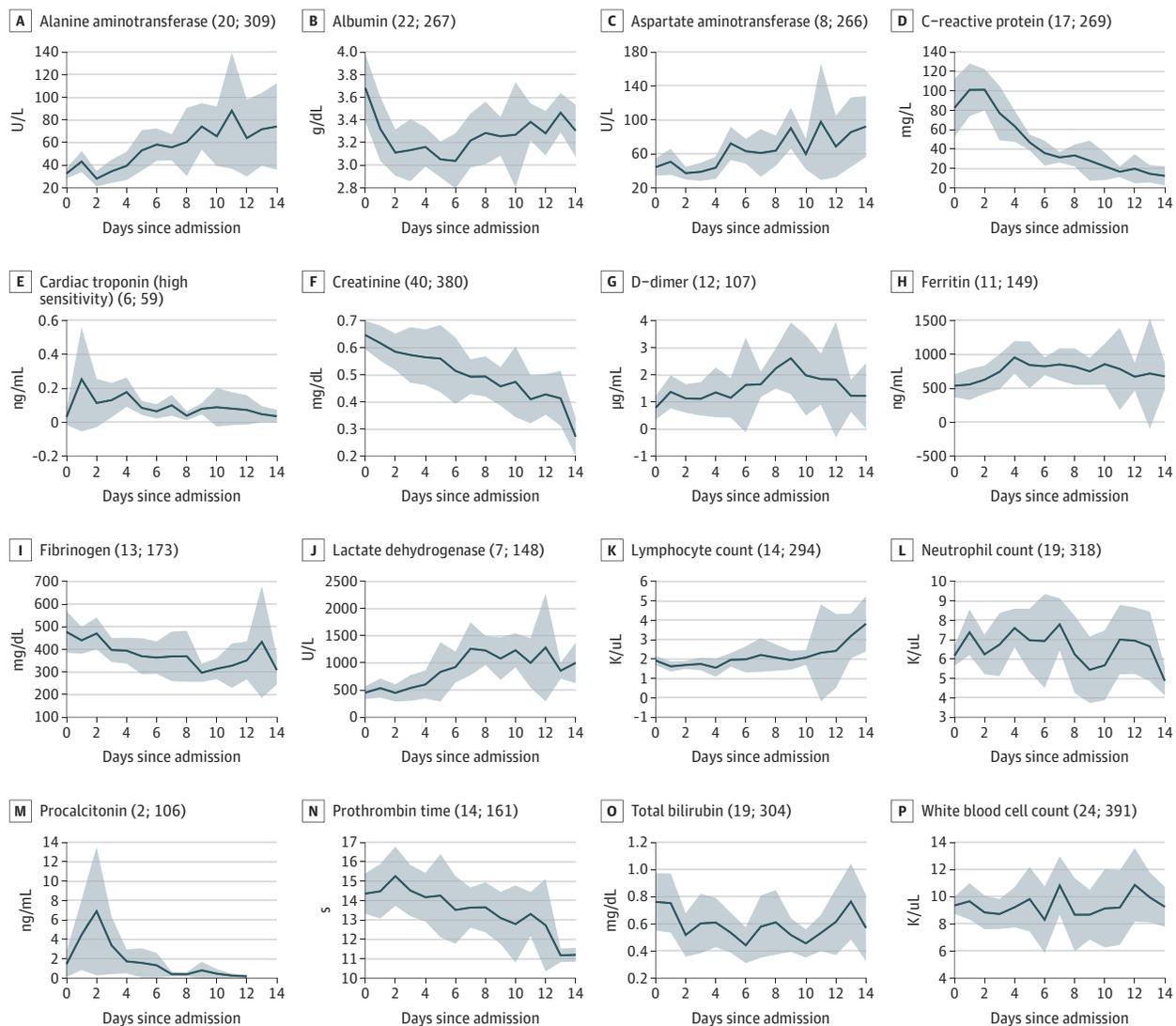
Mean daily values across sites were calculated using random-effects meta-analysis. Values in parenthesis represent the minimum and maximum numbers of patients contributing data on any single day during the 14-day observation period. The shaded areas represent 95% CIs. SI conversion factors: To convert alanine aminotransferase to microkatal per liter, multiply by 0.0167; albumin to g/L, multiply by 10; aspartate aminotransferase to microkatal per liter, multiply by 0.0167; C-reactive protein to milligrams per liter; creatinine to micromoles per liter, multiply by 76.25; ferritin to

microgram per liter, multiply by 1; D-dimer to nanomoles per liter, multiply by 5.476; fibrinogen to grams per liter, multiply by .01; lactate dehydrogenase to microkatal per liter, multiply by 0.0167; lymphocyte count to proportion of 1.0, multiply by 0.01; neutrophil to proportion of 1.0, multiply by 0.01; total bilirubin to micromoles per liter, multiply by 17.104; troponin to milligrams per liter, multiply by 1.0, white blood cell count to proportion of 1.0, multiply by 0.01.

and incomplete recording of certain data elements, 4CE demonstrates how contemporary informatics methods can enable efficient integration and analysis of large volumes of clinical information to build observational data sets across health care systems and countries. A unique feature of this EHR-based network is the rapid onboarding facilitated by the open-source i2b2 and OMOP software platforms, which allowed a total of 96 adult and pediatric hospitals to join during an initial 2-week period. Rapid availability of this type of curated EHR data can support hypothesis generation and prioritization of clinical trials, understanding of the natural course of disease, identification of rare complications and phenotypes, and anticipatory planning by health care institutions around resource requirements and medical supply needs.

Laboratory values were extracted for a core set of tests to support a detailed assessment of the clinical course of patients with COVID-19. Laboratory results are typically not available in administrative or medical claims data sets, which are limited to information on test ordering. We collected daily laboratory values for each day of hospitalization to build trajectories for individual tests. An extensive quality control process was performed to address any mapping errors between sites. However, values should be interpreted with some caution because tests are not obtained consistently on all patients and reflect physician decisions based on the patient's clinical presentation and local health care workflows.[29] Evaluation of the number of patients with available laboratory tests indicated that the number tested dropped quickly during the first week of hospitalization. Nonetheless, our findings were consistent with clinical results presented in case series and meta-analyses, showing generally normal white blood cell counts and abnormally elevated inflammatory markers and coagulation tests.[30-33] Laboratory trajectories also revealed a gradual decline in the value of certain inflammatory markers during hospitalization. Interestingly, several laboratory tests demonstrated increased values beginning around the second week of hospitalization, such as kidney and liver function markers. Additional studies will be needed to determine whether these laboratory tests can predict specific disease trajectories and complications among hospitalized children and youth.

The multinational design of 4CE allows ascertainment of differences in regional management patterns and uptake of therapeutic interventions in children and youth. Early in the pandemic, many agents emerged as candidate therapies for COVID-19, including both repurposed drugs and investigational agents.[17] Observational studies[31,34-38] indicate that many of these drugs were widely used among hospitalized adult patients, although use in children and youth appears to have been lower. This likely reflects the less severe disease course in children and youth and is also consistent with patterns in off-label medication prescribing in pediatric patients, where use in pediatric populations tends to follow adoption in adults.[39] It also relates to the lower number of clinical trials

Table 2. Medication Use

Drug class[a]	Country using the drugs	Health care sites treating at least 3 patients, No, (%) (n = 14)[b]	Patients treated, No.
Repurposed agents			
Aminoquinolines	US	2 (14)	7
Antivirals for HIV and other viral infections	None	0	0
Interferons	None	0	0
Investigational agents			
Remdesivir	UK	1 (7)	6
Adjunctive therapies			
Antithrombotic agents	France, UK, US	8 (57)	137
Diuretics	France, UK, US	8 (57)	97
Interleukin inhibitors	UK, US	3 (21)	12
Angiotensin-converting enzyme inhibitors	US	3 (21)	12
Angiotensin II receptor blockers	None	0	0

[a] Antivirals include lopinavir and ritonavir, darunavir, and ribavirin; aminoquinolines include hydroxychloroquine and chloroquine; interferons include interferon β-1b and peginterferon alfa-2a; and interleukin inhibitors include anakinra, sarilumab, tocilizumab, and siltuximab.

[b] The 27 participating hospitals represent 14 health care sites contributing aggregate data. Hospital-level data are reported according to health care sites, precluding analysis of the number of hospitals providing specific therapies.

performed in pediatric patients to test new therapies, including remdesivir.[2] Monitoring the use of pharmacotherapies in children and youth, including defining regional and country-level differences, will support activities to optimize and standardize care for children and youth with COVID-19 and guide prioritization of research activities to ensure availability of safe and effective pediatric therapies.

An area for further development of 4CE data is in the collection and analysis of race and ethnicity information. During this first phase of data collection, the race variable was limited to standard categories as defined by the US National Institutes of Health and how it is used in many US-based studies.[12] However, this categorization is subject to 2 major limitations for our purposes. First, it combines race and ethnicity in a way that race cannot be reported for Hispanic and Latino individuals. This results in missing race information if a patient is recorded as Hispanic or Latino or missing ethnicity data if race is prioritized in local data collection. Second, this categorization does not lend itself to use in other countries, such as the UK or Singapore, where the primary racial and ethnic categories differ from those in the US. To meaningfully capture race and ethnicity information across countries, country-specific categories must be used. Accurate collection of race and ethnicity information is critical to advancing our understanding of differences in risk factors, infection rates, and health care use that have been reported in prior studies.[15,40,41] In future phases of 4CE, we plan to implement country-specific ontologies for collection of race and ethnicity data.

Limitations

This study has limitations. To enable an international federated network and preserve patient data privacy from each participating site, only aggregate counts were analyzed, limiting the ability to combine values or follow individual patients longitudinally. For example, while we can ascertain mean laboratory values across individual sites and even track these throughout the hospitalization, we cannot link laboratory results to specific patient characteristics. In the next phase of 4CE studies, prespecified analyses will be run within the primary data sets at each of the individual sites before aggregation at the consortium-level, enabling patient-level analyses. Additional limitations relate to the use of observational data, including nonsystematic recording of certain clinical data elements and shifting testing strategies for COVID-19, which may inform characteristics of the study population.[29]

Conclusions

In this study of EHRs of children and youth hospitalized with COVID-19 in 6 countries, we demonstrated country-level variation in trends in COVID-19 hospitalization for children and youth and defined clinical complications and laboratory test abnormalities. Large-scale informatics-based approaches can be applied to complement other methods of disease surveillance and define epidemiological and clinical features of COVID-19 in children and youth. Further study and use of EHR informatics-based efforts may facilitate improved modeling of pediatric COVID-19 trajectories and inform clinical care for pediatric patients.

Corresponding Authors: Paul Avillach, MD, PhD, Department of Biomedical Informatics, Harvard Medical School, 10 Shattuck St, Boston, MA 02115 (paul_avillach@hms.harvard.edu); Florence Bourgeois, MD, MPH, Boston Children's Hospital, 300 Longwood Ave, Boston, MA 02115 (florence.bourgeois@childrens.harvard.edu).

Author Affiliations: Department of Pediatrics, Harvard Medical School, Boston, Massachusetts (Bourgeois);

Computational Health Informatics Program, Boston Children's Hospital, Boston, Massachusetts (Bourgeois, Geva, Mandl, Avillach); Department of Biomedical Informatics, Harvard Medical School, Boston, Massachusetts (Gutiérrez-Sacristán, Keller, Hong, Bonzel, Tan, Gehlenborg, Serret-Larmande, Weber, Kohane, Cai, Avillach); Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts (Liu); Departments of Biomedical Informatics, Pediatrics, Cincinnati Children's Hospital Medical Center, University of Cincinnati, Ohio (Aronow); Institute of Medical Biometry and Statistics, Faculty of Medicine and Medical Center, University of Freiburg, Germany (Boeker, Tippmann); Digital Research, Informatics and Virtual Environments (DRIVE), Great Ormond Street Hospital for Children, London, United Kingdom (Booth, Issitt, Sebire, Spiridou); Department of Health Informatics, Hospital Universitario 12 de Octubre, Madrid, Spain (Cruz Rojo, García Barrio, Jiménez, Balazote); Department of Biomedical Health Informatics and the Department of Pediatrics, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania (Devkota, Taylor); Department of Anesthesiology, Critical Care, and Pain Medicine, Boston Children's Hospital, Boston, Massachusetts (Geva); Department of Learning Health Sciences, University of Michigan, Ann Arbor (Hanauer); Department of Preventive Medicine, Northwestern University, Evanston, Illinois (Hutch, Luo, Mao); Department of Medicine, Massachusetts General Hospital, Boston (Klann); IAM Unit, Bordeaux University Hospital, Bordeaux, France (Moal); Department of Infectious Diseases, Great Ormond Street Hospital for Children, London, United Kingdom (Moshal); Department of Neurology, Massachusetts General Hospital, Boston, Massachusetts (Murphy); Department of Biomedical Informatics, Hôpital Necker-Enfants Malade, Assistance Publique Hôpitaux de Paris, University of Paris, Paris, France (Neuraz); Department of Biomedical informatics, WiSDM, National University Health Systems Singapore, Singapore (Ngiam); Department of Computational Medicine & Bioinformatics, Internal Medicine, Human Genetics, & School of Public Health, University of Michigan, Ann Arbor (Omenn); Department of Internal Medicine, Division of Medical Informatics, University of Kansas Medical Center, Kansas City (Patel); Department of Pediatrics-Section of Nephrology, Brenner Children's Hospital, Wake Forest School of Medicine, Winston Salem, North Carolina (South); Department of Pediatrics, Perelman Medical School at the University of Pennsylvania, Philadelphia (Taylor); Department of Biomedical Informatics, University of Pittsburgh, Pittsburgh, Pennsylvania (Visweswaran).

REFERENCES

1. Shane AL, Sato AI, Kao C, et al. A pediatric infectious disease perspective of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and novel coronavirus disease 2019 (COVID-19) in children. *J Pediatric Infect Dis Soc*. 2020;9(5):596-608. doi:10.1093/jpids/piaa099

2. Hwang TJ, Randolph AG, Bourgeois FT. Inclusion of children in clinical trials of treatments for coronavirus disease 2019 (COVID-19). *JAMA Pediatr*. 2020;174(9):825-826. doi:10.1001/jamapediatrics.2020.1888

3. Jiang L, Tang K, Levin M, et al. COVID-19 and multisystem inflammatory syndrome in children and adolescents. *Lancet Infect Dis*. 2020;20(11):e276-e288. doi:10.1016/S1473-3099(20)30651-4

4. Casey JA, Schwartz BS, Stewart WF, Adler NE. Using electronic health records for population health research: a review of methods and applications. *Annu Rev Public Health*. 2016;37:61-81. doi:10.1146/annurev-publhealth-032315-021353

5. Lang JE, Bunnell HT, Hossain MJ, et al. Being overweight or obese and the development of asthma. *Pediatrics*. 2018;142(6):e20182119. doi:10.1542/peds.2018-2119

6. McMahon AW, Cooper WO, Brown JS, et al. Big data in the assessment of pediatric medication safety. *Pediatrics*. 2020;145(2):e20190562. doi:10.1542/peds.2019-0562

7. Denburg MR, Razzaghi H, Bailey LC, et al. Using electronic health record data to rapidly identify children with glomerular disease for clinical research. *J Am Soc Nephrol*. 2019;30(12):2427-2435. doi:10.1681/ASN.2019040365

8. Fleurence RL, Forrest CB, Shuren J. Strengthening the evidence base for pediatric medical devices using real-world data. *J Pediatr*. 2019;214:209-211. doi:10.1016/j.jpeds.2019.06.060

9. Brat GA, Weber GM, Gehlenborg N, et al. International electronic health record-derived COVID-19 clinical course profiles: the 4CE consortium. *NPJ Digit Med*. 2020;3:109. doi:10.1038/s41746-020-00308-0

10. Klann JG, Weber GM, Estiri H, et al. Validation of a derived international patient severity algorithm to support COVID-19 analytics from electronic health record data. *Health Informatics*. Published online October 14, 2020. doi:10.1101/2020.10.13.20201855

11. Weber GM, Hong C, Palmer NP, et al. International comparisons of harmonized laboratory value trajectories to predict severe COVID-19: leveraging the 4CE collaborative across 342 hospitals and 6 countries: a retrospective cohort study. *bioRxiv*. Published online December 18, 2020. doi:10.1101/2020.12.16.20247684

12. NOT-OD-15-089: racial and ethnic categories and definitions for NIH diversity programs and for other reporting purposes. Accessed November 6, 2020. https://grants.nih.gov/grants/guide/notice-files/not-od-15-089.html

13. Goyal MK, Simpson JN, Boyle MD, et al. Racial and/or ethnic and socioeconomic disparities of SARS-CoV-2 infection among children. *Pediatrics*. 2020;146(4):e2020009951. doi:10.1542/peds.2020-009951

14. Bailey LC, Razzaghi H, Burrows EK, et al. Assessment of 135 794 pediatric patients tested for severe acute respiratory syndrome coronavirus 2 across the United States. *JAMA Pediatr*. 2021;175(2):176-184. doi:10.1001/jamapediatrics.2020.5052

15. Bixler D, Miller AD, Mattison CP, et al; Pediatric Mortality Investigation Team. SARS-CoV-2-associated deaths among persons aged <21 years–United States, February 12-July 31, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69 (37):1324-1329. doi:10.15585/mmwr.mm6937e4

16. Lippi G, Plebani M. Laboratory abnormalities in patients with COVID-2019 infection. *Clin Chem Lab Med*. 2020;58(7):1131-1134. doi:10.1515/cclm-2020-0198

17. Sanders JM, Monogue ML, Jodlowski TZ, Cutrell JB. Pharmacologic treatments for coronavirus disease 2019 (COVID-19): a review. *JAMA*. 2020;323(18):1824-1836. doi:10.1001/jama.2020.6019

18. Viechtbauer W. Conducting meta-analyses in R with the metafor package. *J Stat Softw*. 2010;36(3):1-48. doi:10.18637/jss.v036.i03

19. VanderPlas J, Granger BE, Heer J, et al. Altair: interactive statistical visualizations for python. *J Open Source Softw*. 2018;3(32):1057. doi:10.21105/joss.01057

20. Géodes-Santé publique France. Accessed December 12, 2020. https://geodes.santepubliquefrance.fr/

21. German Society for Pediatric Infectious Diseases. COVID-19 Survey. Accessed December 17, 2020. https://dgpi.de/covid-19-survey-update/

22. Spanish National Epidemiological Surveillance Network. Situación de COVID-19 En España; 2020. Accessed December 17, 2020. https://www.isciii.es/QueHacemos/Servicios/VigilanciaSaludPublicaRENAVE/EnfermedadesTransmisibles/Paginas/-COVID-19.-Informes-previos.aspx

23. Royal College of Paediatrics and Child Health. COVID-19-service evaluation and audit on the care needs of children admitted to hospital (England). Published December 16, 2020. Accessed November 16, 2020. https://www.rcpch.ac.uk/resources/covid-19-service-evaluation-audit-care-needs-children-admitted-hospital-england

24. Healthdata.gov. COVID-19 reported patient impact and hospital capacity by facility. Published December 12, 2020. Accessed December 20, 2020. https://healthdata.gov/dataset/COVID-19-Reported-Patient-Impact-and-Hospital-Capa/6xf2-c3ie

25. 4CE. Consortium for Clinical Characterization of COVID-19 by EHR. Accessed May 24, 2021. https://covidclinical.net/

26. Gaborieau L, Delestrain C, Bensaid P, et al. Epidemiology and clinical presentation of children hospitalized with SARS-CoV-2 infection in suburbs of Paris. *J Clin Med*. 2020;9(7):E2227. doi:10.3390/jcm9072227

27. Swann OV, Holden KA, Turtle L, et al; ISARIC4C Investigators. Clinical characteristics of children and young people admitted to hospital with COVID-19 in United Kingdom: prospective multicentre observational cohort study. *BMJ*. 2020;370:m3249. doi:10.1136/bmj.m3249

28. Schulz WL, Kvedar JC, Krumholz HM. Agile analytics to support rapid knowledge pipelines. *NPJ Digit Med*. 2020;3:108. doi:10.1038/s41746-020-00309-z

29. Cohen JB, D'Agostino McGowan L, Jensen ET, Rigdon J, South AM. Evaluating sources of bias in observational studies of angiotensin-converting enzyme inhibitor/angiotensin II receptor blocker use during COVID-19: beyond confounding. *J Hypertens*. 2021;39(4):795-805. doi:10.1097/HJH.0000000000002706

30. Henry BM, Benoit SW, de Oliveira MHS, et al. Laboratory abnormalities in children with mild and severe coronavirus disease 2019 (COVID-19): a pooled analysis and review. *Clin Biochem*. 2020;81:1-8. doi:10.1016/j.clinbiochem.2020.05.012

31. Zachariah P, Johnson CL, Halabi KC, et al; Columbia Pediatric COVID-19 Management Group. Epidemiology, clinical features, and disease severity in patients with coronavirus disease 2019 (COVID-19) in a children's hospital in New York City, New York. *JAMA Pediatr*. 2020;174(10):e202430. doi:10.1001/jamapediatrics.2020.2430

32. Kainth MK, Goenka PK, Williamson KA, et al; Northwell Health COVID-19 Research Consortium. Early experience of COVID-19 in a US children's hospital. *Pediatrics*. 2020;146(4):e2020003186. doi:10.1542/peds.2020-003186

33. de Souza TH, Nadal JA, Nogueira RJN, Pereira RM, Brandão MB. Clinical manifestations of children with COVID-19: a systematic review. *Pediatr Pulmonol*. 2020;55(8):1892-1899. doi:10.1002/ppul.24885

34. Argenziano MG, Bruce SL, Slater CL, et al. Characterization and clinical course of 1000 patients with coronavirus disease 2019 in New York: retrospective case series. *BMJ*. 2020;369:m1996. doi:10.1136/bmj.m1996

35. Lecronier M, Beurton A, Burrel S, et al. Comparison of hydroxychloroquine, lopinavir/ritonavir, and standard of care in critically ill patients with SARS-CoV-2 pneumonia: an opportunistic retrospective analysis. *Crit Care*. 2020; 24(1):418. doi:10.1186/s13054-020-03117-9

36. Gupta S, Wang W, Hayek SS, et al; STOP-COVID Investigators. Association between early treatment with tocilizumab and mortality among critically ill patients with COVID-19. *JAMA Intern Med*. 2021;181(1):41-51. doi:10.1001/jamainternmed.2020.6252

37. Otto WR, Geoghegan S, Posch LC, et al. The epidemiology of severe acute respiratory syndrome coronavirus 2 in a pediatric healthcare network in the United States. *J Pediatric Infect Dis Soc*. 2020;9(5):523-529. doi:10.1093/jpids/piaa074

38. Chao JY, Derespina KR, Herold BC, et al. Clinical characteristics and outcomes of hospitalized and critically ill children and adolescents with coronavirus disease 2019 at a tertiary care medical center in New York City. *J Pediatr*. 2020;223:14-19.e2. doi:10.1016/j.jpeds.2020.05.006

39. Lindkvist J, Airaksinen M, Kaukonen AM, Klaukka T, Hoppu K. Evolution of paediatric off-label use after new significant medicines become available for adults: a study on triptans in Finnish children 1994-2007. *Br J Clin Pharmacol*. 2011;71(6):929-935. doi:10.1111/j.1365-2125.2010.03881.x

40. Kabarriti R, Brodin NP, Maron MI, et al. Association of race and ethnicity with comorbidities and survival among patients with COVID-19 at an urban medical center in New York. *JAMA Netw Open*. 2020;3(9):e2019795. doi:10.1001/jamanetworkopen.2020.19795

41. Ogedegbe G, Ravenell J, Adhikari S, et al. Assessment of racial/ethnic disparities in hospitalization and mortality in patients with COVID-19 in New York City. *JAMA Netw Open*. 2020;3(12):e2026881. doi:10.1001/jamanetworkopen.2020.26881

SUPPLEMENT 1.
eTable. Contributing sites
eFigure 1. Federated data collection across participating sites
eFigure 2. Forest plots for laboratory values at admission
eFigure 3. Sample size plot
eMethods. Calculation of temporal trends for laboratory values

SUPPLEMENT 2.
Nonauthor Collaborators. The Consortium for Clinical Characterization of COVID-19 by EHR (4CE) coordinators and investigators