scientific reports

OPEN

A retrospective cohort study of 12,306 pediatric COVID-19 patients in the United States

Vibhu Parcha[1], Katherine S. Booker[2,3], Rajat Kalra[4], Seth Kuranz[5], Lorenzo Berra[6], Garima Arora[1] & Pankaj Arora[1,7✉]

Children and adolescents account for ~ 13% of total COVID-19 cases in the United States. However, little is known about the nature of the illness in children. The reopening of schools underlines the importance of understanding the epidemiology of pediatric COVID-19 infections. We sought to assess the clinical characteristics and outcomes in pediatric COVID-19 patients. We conducted a retrospective cross-sectional analysis of pediatric patients diagnosed with COVID-19 from healthcare organizations in the United States. The study outcomes (hospitalization, mechanical ventilation, critical care) were assessed using logistic regression. The subgroups of sex and race were compared after propensity score matching. Among 12,306 children with lab-confirmed COVID-19, 16.5% presented with respiratory symptoms (cough, dyspnea), 13.9% had gastrointestinal symptoms (nausea, vomiting, diarrhea, abdominal pain), 8.1% had dermatological symptoms (rash), 4.8% had neurological (headache), and 18.8% had other non-specific symptoms (fever, malaise, myalgia, arthralgia and disturbances of smell or taste). In the study cohort, the hospitalization frequency was 5.3%, with 17.6% needing critical care services and 4.1% requiring mechanical ventilation. Following propensity score matching, the risk of all outcomes was similar between males and females. Following propensity score matching, the risk of hospitalization was greater in non-Hispanic Black (RR 1.97 [95% CI 1.49–2.61]) and Hispanic children (RR 1.31 [95% CI 1.03–1.78]) compared with non-Hispanic Whites. In the pediatric population infected with COVID-19, a substantial proportion were hospitalized due to the illness and developed adverse clinical outcomes.

Over 4.2 million children in the United States have tested positive for coronavirus disease-2019 (COVID-19) since the onset of the pandemic[1,2]. In comparison with adults, preliminary reports suggest that children (< 18 years of age) have relatively lower odds of adverse clinical outcomes associated with COVID-19[3–8]. The lower observed prevalence of COVID-19 in the pediatric age-group worldwide is partially attributed to wide-spread school closures in response to the pandemic[7,9,10]. Furthermore, challenges in the adequate screening and testing of children, especially those who are asymptomatic or minimally symptomatic, may have also contributed to the underreporting of COVID-19 in children. The cautious reopening of schools in the United States and other countries has occurred in the backdrop of an increased possibility of community transmission of COVID-19 among children in schools[7,8,10]. Thus, it is important to characterize the demographic, clinical characteristics, and outcomes in children infected with COVID-19. There are limited data, especially from the United States, describing the demographics, clinical characteristics, and outcomes of lab-confirmed COVID-19 children[3–8]. We present the findings of an investigation evaluating the clinical characteristics, comorbidities, and complications in 12,306 lab-confirmed COVID-19 patients from a multicenter federated healthcare network electronic health record database.

[1]Division of Cardiovascular Disease, University of Alabama at Birmingham, 1670 University Boulevard, Volker Hall B140, Birmingham, AL 35294-0019, USA. [2]Department of Internal Medicine, Abbott Northwestern Hospital, Minneapolis, MN, USA. [3]Division of Hospital Medicine, Children's Minnesota, Minneapolis, MN, USA. [4]Cardiovascular Division, University of Minnesota, Minneapolis, MN, USA. [5]TriNetX, Inc., Cambridge, MA, USA. [6]Anesthesia & Critical Care, Pulmonary Medicine, Massachusetts General Hospital, Boston, MA, USA. [7]Section of Cardiology, Birmingham Veterans Affairs Medical Center, Birmingham, AL, USA. ✉email: parora@uabmc.edu
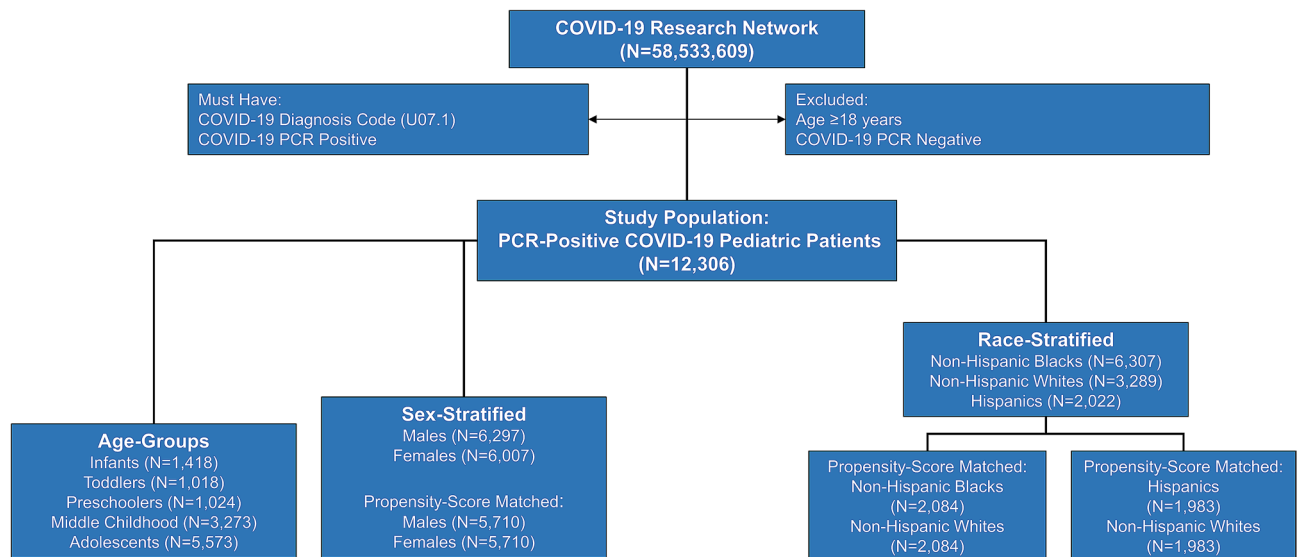
Figure 1. Derivation of study population.

Results

We identified 12,306 children with lab-confirmed COVID-19 in our cohort from 33 healthcare organizations, which are part of the TriNetX Research Network. Figure 1 describes the derivation of the study population. The geographic distribution of the cohort contributing to this analysis was composed of 43.0% patients (5289) from the Southern region, 22.9% (2814) patients from the Western region, 22.2% (2734) patients from the Midwestern region, and 7.2% (886) patients from the Northeastern region. The clinical characteristics of the overall study population stratified by hospitalization status are described in Table 1. At the time of presentation, the children who were hospitalized had a greater prevalence of fever, respiratory, gastrointestinal symptoms, and a relatively higher prevalence comorbidities. The frequency of presenting symptoms is shown in Fig. 2. Children who were hospitalized had relatively lower hemoglobin and neutrophil count. Among the study population, only 25.1% of children had at least one of the typical symptoms (fever, cough, or shortness of breath), and 9.9 of children had at least two typical symptoms. Nearly three-fourths (74.9%) of the children did not have any of the typical COVID-19 symptoms. In the study population, 16.5% presented with respiratory symptoms (cough, dyspnea), 13.9% had gastrointestinal symptoms (nausea, vomiting, diarrhea, abdominal pain), 8.1% had dermatological symptoms (rash), 4.8% had neurological (headache), and 18.8% had other non-specific symptoms (fever, malaise, myalgia, arthralgia and disturbances of smell or taste). The lab parameters of the hospitalized patients are described in Table 2. In the overall cohort, the frequency of hospitalization was 5.5%. Among those who were hospitalized (N = 672), 17.6% required critical care services (N = 118), and 4.1% required mechanical ventilation (N = 38). There were ≤ 10 deaths in the study population. Among those without any comorbidities (N = 8,297), the frequency of hospitalization was 3.5%, 0.4% required critical care, and 0.3% required mechanical ventilation.

In our sex-stratified comparison of the study population, 5,710 pairs of male and female patients were identified after propensity-score matching. The pre and post matching propensity-score density for the subgroup comparisons are depicted in Supplementary Figs. 2–4. In the matched pairs, the risk of hospitalization was similar (RR 0.94 [95% CI 0.80–1.12]) between males (4.5%) and females (4.7%). Compared with females (0.7%), the risk ratio of requiring critical care among males (0.9%) was 0.87 (95% CI 0.55–1.12). The incidence of mechanical ventilation (RR 0.93 [95% CI 0.44–1.97]) was similar between males (0.2%) and females (0.3%).

The age-stratified presenting symptoms of the patients are described in Fig. 3. The prevalence of fever and rash was lower in adolescents compared with younger children and infants. The frequency of hospitalization, critical care requirement, and mechanical ventilation stratified by age groups is depicted in Supplementary Table 1.

After propensity-score matching, we identified 2084 pairs of non-Hispanic White and non-Hispanic Black children. The risk of hospitalization was higher in non-Hispanic Black children with COVID-19 (6.5%) compared with non-Hispanic White (3.3%) children (RR 1.97 [95% CI 1.49–2.61]). The risk of requiring critical care (RR 1.71 [95% CI 0.89–3.31]) and mechanical ventilation (RR 1.00 [95% CI 0.42–2.40]) was similar in non-Hispanic Blacks (1.2% and 0.5%) and non-Hispanic White (0.7% and 0.5%) pediatric patients. We identified 1983 propensity-score-matched pairs of Hispanic and non-Hispanic White patients. The risk of hospitalization (RR 1.31 [95% CI 1.03–1.78]) was higher in Hispanics (4.6%) compared with non-Hispanic Whites (3.5%). The risk of requiring critical care (RR 1.18 [95% CI 0.53–2.63]) and mechanical ventilation (RR 1.00 [95% CI 0.41–2.39]) was similar in Hispanic (0.7% and 0.5%) and non-Hispanic White (0.5% and 0.5%) pediatric patients. The frequency of study outcomes in non-Hispanic Asians and other racial/ethnic groups is demonstrated in Supplementary Table 2.

The descriptive trends in the case count showed a consistent increase since April 1, 2020. However, the number of hospitalization peaked in July 2020 in our cohort and was followed by a decline subsequently (Supplementary Fig. 5).

Characteristics	Non-hospitalized cohort (total N = 11,634)		Hospitalized cohort (total N = 672)		P-value	Standardized mean difference
	Patient count	Mean ± SD or proportion	Patient count	Mean ± SD or proportion		
Age (years)	11,634	9.4 ± 5.6	672	9.0 ± 6.2	0.13	0.06
Age groups					< 0.001	0.53
Infant (0–1 years)	1345	11.6%	73	10.9%		
Toddlers (1–3 years)	933	8.0%	85	12.6%		
Preschoolers (3–5 years)	959	8.2%	65	9.7%		
Middle childhood (6–11 years)	3146	27.0%	127	18.9%		
Adolescents (≥ 12 years)	5256	45.1%	317	47.1%		
Male	5954	51.2%	343	51.0%	0.94	0.002
Female	5678	48.8%	329	49.0%		
Race/ethnicity						
Non-Hispanic Whites	6064	52.1%	243	36.2%	< 0.001	0.18
Non-Hispanic Blacks	3066	26.4%	223	33.2%	< 0.001	0.21
Hispanic	1889	16.2%	133	19.8%	< 0.001	0.16
Non-Hispanic Asians	273	2.3%	19	2.8%	0.43	0.03
Other	342	2.9%	54	8.0%	0.04	0.08
Comorbidities						
Cardiovascular	94	0.8%	102	15.2%	< 0.001	0.54
Gastrointestinal	334	2.9%	121	18.0%	< 0.001	0.51
Hematologic or immunologic	57	0.5%	45	6.7%	< 0.001	0.30
Malignancy	51	0.4%	34	5.1%	< 0.001	0.29
Metabolic	79	0.8%	69	10.3%	< 0.001	0.38
Neurological and Neuromuscular	252	2.2%	92	13.7%	< 0.001	0.44
Congenital or genetic defects	127	1.1%	191	28.4%	< 0.001	0.83
Renal and urologic	175	1.5%	87	13.0%	< 0.001	0.45
Respiratory	2,310	19.9%	189	28.1%	< 0.001	0.20

Table 1. Baseline characteristics of the overall cohort. *SD* standard deviation.

Discussion

In this study of children with COVID-19, we observed a high prevalence of non-specific symptoms at presentation, with frequent multi-organ involvement. In our study cohort, ~ 5% were hospitalized, and among those who were hospitalized, ~ 18% required critical care, and ~ 4% needed mechanical ventilation. The clinical outcomes were similar across subgroups of sex. Non-Hispanic Black and Hispanic children with COVID-19 had a higher risk of hospitalization when compared with non-Hispanic White children. The temporal trend of the cases and hospitalization was similar to the nationwide population trends for COVID-19 cases.

We confirm the findings of prior reports from smaller populations describing the relatively milder clinical course and a relatively lower incidence of adverse clinical outcomes among children compared to adults[8,11–15]. The findings from our reports describe the wide spectrum of illness seen in children with COVID-19 across the demographic and age subgroups. The recognition of these clinical characteristics is important for the early identification and care of children with COVID-19. We observed a higher frequency of hospitalization in non-Hispanic Blacks and Hispanics. This is concordant with the recent findings reported by the Centers for Disease Control and Prevention (CDC) from the hospitals participating in the COVID-NET database[16]. The observed racial differences may be due to greater indirect viral exposure to children from racial/ethnic minorities due to the various socioeconomic impediments to implementation of infection control measures[17–24]. These racial disparities have been previously noted during the H1N1 pandemic[23,25,26]. These racial differences in COVID-19 burden are also evident in the current data from the pediatric intensive care units in the United States[27]. We observed a similar trend in COVID-19 cases and hospitalization in our database as being reported by larger tracking databases such as that by the COVKID project and the CDC's COVID-NET[2,28,29].

We noted that the majority of our study population were not recorded to have typical symptoms (i.e., fever, cough, or dyspnea). Prior investigations in different populations, in diverse settings, and with varying age distributions have reported that up to 40–70% of pediatric patients may present with fever and respiratory symptoms[31,32]. In contrast, some investigations have also reported a high prevalence (up to ~ 50%) of asymptomatic or mild COVID-19 infections in children[33,34]. A previous study in the United States[32] had noted that testing indications were unclear in nearly 50% of the children, which may contribute to the low prevalence of typical
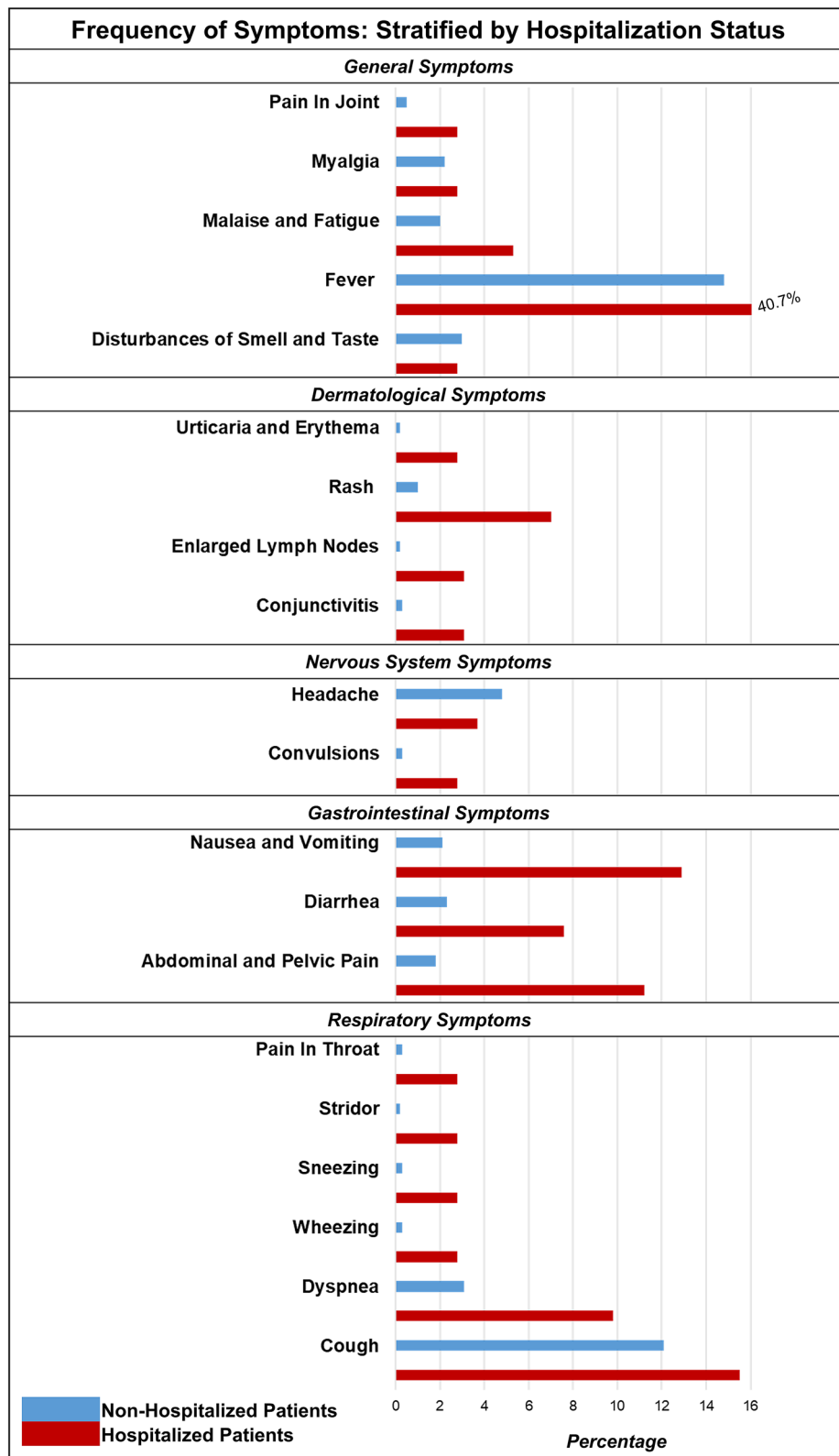
Figure 2. Frequency of symptoms in pediatric COVID-19 patients: stratified by hospitalization.

symptoms at the time of diagnosis. There may also be additional factors contributing to COVID-19 testing among children, such as exposure to SARS-CoV-2 infected individuals, parental vigilance, and asymptomatic screening

Characteristics	Hospitalized cohort (total N = 672)	
	Patient count	Mean ± SD
Hematological parameters		
Leukocytes (Cells × 10^3/μL)	446	8.8 ± 5.0
Neutrophils (Cells × 10^3/μL)	336	5.8 ± 3.9
Lymphocytes (per 100 leukocytes)	339	28.2 ± 9.4
Monocytes (per 100 leukocytes)	359	7.7 ± 4.6
Hemoglobin (g/dL)	462	11.9 ± 2.2
Platelets (Cells × 10^3/μL)	446	294 ± 116
Liver function		
Alanine aminotransferase (IU/L)	478	42.6 ± 8.8
Albumin (g/dL)	298	3.9 ± 0.7
Alkaline phosphatase (IU/L)	309	179.0 ± 105.0
Aspartate aminotransferase (IU/L)	315	85.3 ± 57.0
Bilirubin total (mg/dL)	306	0.5 ± 0.3
Protein (g/dL)	299	6.7 ± 1.2
Renal function		
Creatinine (mg/dL)	384	0.7 ± 0.4
Blood urea nitrogen (mg/dL)	308	11.5 ± 5.9
Electrolytes		
Calcium (mg/dL)	371	9.2 ± 0.6
Potassium (mEq/L)	348	4.2 ± 0.7
Sodium (mEq/L)	379	139.0 ± 4.0
Coagulation function		
Activated partial thromboplastin time (s)	139	32.2 ± 10.2
Prothrombin time (s)	149	14.9 ± 6.0
INR	152	1.2 ± 0.6

Table 2. Laboratory measures of the hospitalized patients. *SD* standard deviation.
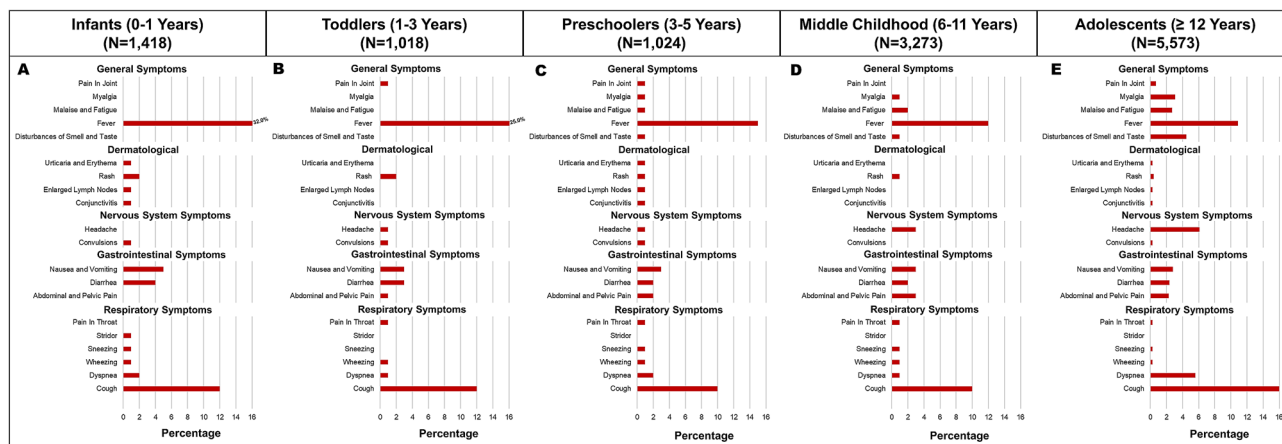

Figure 3. Frequency of symptoms in pediatric COVID-19 patients: stratified by age-groups.

for travel or surgery[8,30,33–39]. However, similar to our findings, nearly all prior studies have found a relatively high proportion of non-specific signs and symptoms prompting testing among pediatric COVID-19 patients, including lethargy, malaise, myalgia, sore throat, runny nose, sneezing, gastrointestinal symptoms, and fatigue[8,30–39]. The relatively lower rates of typical symptoms noted in our study compared with other studies may also be due to the incomplete reporting of symptoms in the electronic health records, difficulty in eliciting symptomology in pre-verbal pediatric patients, relatively higher proportion of non-typical symptoms, geographic differences in the extent of spread of COVID-19[40], and differences in the local screening and testing approaches. Further evaluation of the clinical presentation of COVID-19 among pediatric populations is needed to adequately target our screening and testing approaches.

The age-related differences in COVID-19 population prevalence[41] and associated clinical outcomes may be a result of multiple factors[3]. Poor clinical outcomes among adult COVID-19 patients are associated with a

higher comorbidity burden[42–45]. Children usually do not exhibit multiple comorbidities till a later age, and this may contribute to the lower rates of adverse clinical outcomes in children compared with adults. Pre-existing empiric immunity as a result of frequent seasonal human coronavirus infections has also been hypothesized to contribute to the lower SARS-CoV-2 infection rate among children and adolescents compared with adults[46]. In the absence of patient-level data about prior infections, the role of empiric immunity in the prevention of infection and its clinical manifestation requires further investigation. There may be other factors that may contribute to the observed lower risk in children compared with adults, such as age-dependent expression of ACE2 receptor (SARS-CoV-2 binding receptor) and androgen levels[47–49]. We also observed relatively lower neutrophil levels in the hospitalized pediatric COVID-19 patients indicating a role of age-related neutrophil recruitment in the mild manifestation of the illness in children, as reported previously[50]. The lower disease prevalence and severity in children may be due to both having lower susceptibility to COVID-19 infection and a lower likelihood of showing symptoms[3,4].

Our findings have public health implications. The initial risk of infection transmission among children may have been limited to some extent by the early closure of schools, colleges, and universities[9,10]. The public health impact of the school closure and reopening is not completely understood[9,10]. While the presented data does not capture the potential public health impact of the various measures, these data may help to address the lacunae of clinical evidence around the patient characteristics of children with COVID-19. We observed a higher prevalence of comorbidities among those who were hospitalized following COVID-19. Identifying children who are at greater risk of complications[51–53] may serve to create tailored strategies to screen them aggressively. Overall, our findings suggest that children and adolescents may have a milder course of illness compared with adults with COVID-19[54]. Given the high prevalence of non-specific signs and symptoms and the fact that the majority of the patients lacked typical[55] symptoms in our investigation, increased vigilance, innovative screening, and frequent testing is required among school-going children and their immediate contacts. Routine screening tools and procedures such as daily temperature checks in school may be less effective. Our study findings may guide the resource utilization and mitigation efforts by local and federal health authorities, especially in areas with high COVID-19 incidence and prevalence. Innovative approaches, such as sentinel surveillance, random testing of children and the teachers, prioritizing children from high-risk households for COVID-19 testing, and providing education and training on the appropriate use of non-invasive pulse oximeters, may yield additional benefits and help mitigate the spread of COVID-19 among children. Implementation of these strategies may need to be enhanced among children from racial/ethnic minorities to curtail the existing COVID-19 related health disparities.

There are several limitations to our study. The patient exposure and outcomes are defined using administrative codes, which may be subject to coding errors[56,57]. Similarly, it is difficult to parse out the severity of the clinical outcomes[58–63]. Importantly, there may be significant selection bias in the children who were tested based on the indication for obtaining the tests, availability of tests, and access to testing locations and hence the eventual inclusion of the children in the study population. There were also periods early in the pandemic where testing was primarily advised for children whose clinical symptoms were thought to represent a high likelihood of COVID-19. Additionally, the information presented in our investigation is accrued from the structured data recorded in the electronic documentation. Thus, the data may be subject to inaccuracies and incomplete reporting. This may be especially important in the adequate documentation of clinical symptoms in the patient records. Furthermore, laboratory markers may not have been collected in all patients, and there may be an indication bias in the reporting of those results. Additionally, the ability to elicit symptomology is naturally limited by the nature of pediatric medicine. There may also be under-reporting of comorbidities in the administrative datasets[56,57,64–66]. We included all patients with race/ethnicity data available. While this may contribute to some degree of selection bias, it allows for accurate assessment of race-stratified outcomes. We used all-cause hospitalization as the study outcome, which makes it difficult to ascertain whether the hospitalization was due to COVID-19 or another cause. Due to lack of availability of the raw dataset, we were unable to compute the time to hospitalization and the duration of hospitalization from these data. Specific manifestations, such as the multi-systemic inflammatory syndrome[12] in COVID-19, may not have a uniform description in electronic documentation. Thus, we could not evaluate the prevalence of this important sequela of COVID-19. Our study is also limited by the ability to detect the transmission potential of the diagnosed patients. Due to the obfuscation of counts of less than 10 for privacy concerns, we are unable to report the exact number of deaths in the overall population and subgroups.

In summary, children infected with COVID-19 present with a broad spectrum of non-specific symptoms across the age groups. Children with COVID-19 can develop severe illness requiring hospitalization and critical care, but the rates of severe illness and death are relatively low.

Methods

Data source. The TriNetX (Cambridge, MA) COVID-19 Research Network database was used for this study[58–63]. This research network database is a federated health research network database that incorporates and integrates the electronic health records from the participating healthcare organizations, which includes nearly 59 million patients. The research network database integrates cloud-based HIPAA-compliant real-time aggregate patient-level data from the electronic health records, which includes diagnoses, procedures, medication use, and clinical laboratory values from the contributing organizations[58–63]. The participating organizations contribute data from inpatient, outpatient, and specialty services. The TriNetX database integrates data from all the participating organizations after clearance through local data warehouses and research data repositories. To ensure patient privacy, the stored and transmitted data are de-identified at the patient, and organization level. Structured data recorded in the electronic health records are assimilated into the database after mapping the data to standard and controlled clinical terms. A rigorous data quality assessment is done to exclude records that do

not meet quality standards and basic formatting requirements for adequate data representation[63]. The referential integrity is maintained to ensure comparison of data across several databases. Moreover, TriNetX software also ensures data validity by regularly monitoring the temporal trend of data volume[58–61,63]. Further details about the database are provided in the Supplementary Methods[63]. This study was deemed to be non-human subjects research by the University of Alabama at Birmingham Institutional Reviewer Board.

Standardized coding of the database.   The TriNetX database captures diagnoses using the International Statistical Classification of Diseases and Related Health Problems Clinical Modifications 10th Edition (ICD-10-CM). The database records the procedures using standardized Current Procedural Terminology (CPT) codes. The clinical lab parameters are captured using the standard Logical Observation Identifiers Names and Codes (LOINC). The individual codes chosen for defining the COVID-19 patients and for the definition of study outcomes are described in Supplementary Table 4[58–61].

Study population.   The COVID-19 research network database was queried to identify individuals < 18 years of age and diagnosed with lab-confirmed COVID-19 between April 1, 2020 to October 31, 2020. Patients with the relevant International Statistical Classification of Diseases and Related Health Problems (ICD) codes (U07.1)[67] or positive lab result (LOINC Codes: 94309-2, 94315-9, 94316-7, 94500-6, 94533-7, 94534-5, 94502-2, 94599-2, 9088, 40458-1) were included in the study based on the World Health Organization and Centers for Disease Control and Prevention Guidelines. We included only individuals with lab-confirmed (using positive polymerase-chain-reaction) COVID-19 results to ensure that we do not include children with suspected COVID-19, i.e., patients under investigation. We further stratified the population on the basis of age (infants [0–1 year], toddlers [1–3 years], preschoolers [3–5 years], middle childhood [6–11 years], and adolescents [12–17 years]), and sex (males and females). The population was also stratified on the basis of race/ethnicity (non-Hispanic White, non-Hispanic Black, Hispanic, non-Hispanic Asian, Others). The study population was further stratified patients based on the month of diagnosis (April, May, June, July, August, September, October) to identify the temporal trend in the cases and outcomes.

Measures and outcomes.   We identified the baseline patient characteristics, including past medical history, presenting symptoms, medications, and lab parameters. We identified the clinical features and laboratory values for lab parameters, which were identified from within the last 1 month up to the index event in the hospitalized cohort. Typical symptoms were defined as having any of the three symptoms: fever, cough, and shortness of breath, as defined by the CDC[55]. The lab-confirmed diagnosis of COVID-19 was defined as the index event. The standardized ICD-10 diagnosis codes were used to identify the history of existing medical conditions. The main study outcome was the frequency of all-cause hospitalization within 30-days of testing positive in children with COVID-19 and in the abovementioned sub-groups. The additional study outcomes included mechanical ventilation and the requirement for critical care. The administrative diagnosis and procedural codes were used for the identification of the aforementioned study outcomes (Supplementary Tables 4–6).

Statistical analyses.   We summarized the baseline characteristics as mean ± standard deviation for continuous data and as numbers and percentages for categorical data. The baseline characteristics were compared using descriptive statistics with the continuous data were compared using independent sample t-test, and categorical data were compared using the z-score. The study outcomes were compared in age and sex subgroups. All primary and secondary outcomes were reported in the overall populations and the sub-groups of age, sex, and race/ethnicity. Logistic regression was applied to obtain a propensity score for each patient using logistic regression implemented by the function LogisticRegression of the scikit-learn package in Python version 3.7[63,68]. The output was verified by repeating the propensity scoring in R version 3.4.4. Subsequently, a 1:1 matching was done using greedy nearest neighbor matching with a caliper of 0.1 pooled standard deviation[63,69]. The propensity score-matched populations were matched for age, sex, race, and comorbidities (cardiovascular, respiratory, gastrointestinal, malignancy, metabolic, hematological or immunological, neurological and neuromuscular, congenital or genetic defects, renal or urological) (Supplementary Table 6)[70]. For the protection of inadvertent disclosure of protected health information, patient counts for demographics, clinical characteristics, and outcomes if less than is reported as ≤ 10. We report the comparative risk of the study outcomes as risk ratios with 95% confidence intervals. The two-tailed type I error of 0.05 was deemed to be statistically significant. The cloud-based TriNetX analytics platform, which utilizes a combination of JAVA, R, and Python, was used for all analyses[58–61,63].

Data availability
The data from the TriNetX COVID-19 Research database is available to member healthcare organizations through the online cloud-based TriNetX research platform available at https://www.trinetx.com/. The aggregate patient-level is integrated from the electronic health records of the member healthcare organizations, with data available for download by request at the participating institutions.

References
1. American Academy of Pediatrics and Children's Hospital Association. Children and COVID-19: State Data Report (Version 9/17/20). https://services.aap.org/en/pages/2019-novel-coronavirus-covid-19-infections/children-and-covid-19-state-level-data-report/. Accessed 26 Sep 2020.

2. The COVKID Project (Coronavirus in Kids Tracking and Education Project). https://www.covkidproject.org/. Accessed 8 April 2021.
3. Davies, N. G. *et al.* Age-dependent effects in the transmission and control of COVID-19 epidemics. *Nat. Med.* https://doi.org/10.1038/s41591-020-0962-9 (2020).
4. Stringhini, S. *et al.* Seroprevalence of anti-SARS-CoV-2 IgG antibodies in Geneva, Switzerland (SEROCoV-POP): A population-based study. *Lancet* **396**, 313–319. https://doi.org/10.1016/S0140-6736(20)31304-0 (2020).
5. Posfay-Barbe, K. M. *et al.* COVID-19 in children and the dynamics of infection in families. *Pediatrics* https://doi.org/10.1542/peds.2020-1576 (2020).
6. CDC COVID Data Tracker. https://www.cdc.gov/covid-data-tracker/. Accessed 3 Aug 2020.
7. Armitage, R. & Nellums, L. B. Considering inequalities in the school closure response to COVID-19. *Lancet Glob. Health* **8**, e644. https://doi.org/10.1016/S2214-109X(20)30116-9 (2020).
8. Dong, Y. *et al.* Epidemiology of COVID-19 among children in China. *Pediatrics* https://doi.org/10.1542/peds.2020-0702 (2020).
9. Donohue, J. M. & Miller, E. COVID-19 and school closures. *JAMA* https://doi.org/10.1001/jama.2020.13092 (2020).
10. Auger, K. A. *et al.* Association between statewide school closure and COVID-19 incidence and mortality in the US. *JAMA* https://doi.org/10.1001/jama.2020.14348 (2020).
11. Singh-Grewal, D. *et al.* Update on the COVID-19-associated inflammatory syndrome in children and adolescents; paediatric inflammatory multisystem syndrome-temporally associated with SARS-CoV-2. *J. Paediatr. Child Health* https://doi.org/10.1111/jpc.15049 (2020).
12. Feldstein, L. R. *et al.* Multisystem inflammatory syndrome in U.S. children and adolescents. *N. Engl. J. Med.* **383**, 334–346. https://doi.org/10.1056/NEJMoa2021680 (2020).
13. Cheung, E. W. *et al.* Multisystem inflammatory syndrome related to COVID-19 in previously healthy children and adolescents in New York City. *JAMA* https://doi.org/10.1001/jama.2020.10374 (2020).
14. Liu, W. *et al.* Detection of Covid-19 in children in early January 2020 in Wuhan, China. *N. Engl. J. Med.* **382**, 1370–1371. https://doi.org/10.1056/NEJMc2003717 (2020).
15. Parri, N., Lenge, M., Buonsenso, D. & Coronavirus Infection in Pediatric Emergency Departments Research, G. Children with Covid-19 in Pediatric Emergency Departments in Italy. *N. Engl. J. Med.* **383**, 187–190. https://doi.org/10.1056/NEJMc2007617 (2020).
16. Kim, L., Whitaker, M., O'Halloran, A., *et al.* Hospitalization rates and characteristics of children aged <18 years hospitalized with laboratory-confirmed COVID-19—COVID-NET, 14 States, March 1–July 25, 2020. *MMWR Morb. Mortal Wkly. Rep.* (**ePub: 7 August 2020**).
17. Chowkwanyun, M. & Reed, A. L. Jr. Racial health disparities and Covid-19—Caution and context. *N. Engl. J. Med.* https://doi.org/10.1056/NEJMp2012910 (2020).
18. Yancy, C. W. COVID-19 and African Americans. *JAMA* https://doi.org/10.1001/jama.2020.6548 (2020).
19. Garg, S. *et al.* Hospitalization Rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019—COVID-NET, 14 States, March 1–30, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 458–464. https://doi.org/10.15585/mmwr.mm6915e3 (2020).
20. Centers for Disease Control and Prevention. COVID-19 in Racial and Ethnic Minority Groups. https://www.cdc.gov/coronavirus/2019-ncov/need-extra-precautions/racial-ethnic-minorities.html. Accessed 5 May 2020.
21. Jackson, S. A., Anderson, R. T., Johnson, N. J. & Sorlie, P. D. The relation of residential segregation to all-cause mortality: A study in black and white. *Am. J. Public Health* **90**, 615–617. https://doi.org/10.2105/ajph.90.4.615 (2000).
22. Hall, W. J. *et al.* Implicit racial/ethnic bias among health care professionals and its influence on health care outcomes: A systematic review. *Am. J. Public Health* **105**, e60-76. https://doi.org/10.2105/AJPH.2015.302903 (2015).
23. Sachedina, N. & Donaldson, L. J. Paediatric mortality related to pandemic influenza A H1N1 infection in England: An observational population-based study. *Lancet* **376**, 1846–1852. https://doi.org/10.1016/S0140-6736(10)61195-6 (2010).
24. Parcha, V. *et al.* Geographic variation of racial disparities in health and COVID-19 mortality. *Mayo Clin. Proc. Innov. Qual. Outcomes* https://doi.org/10.1016/j.mayocpiqo.2020.09.005 (2020).
25. Quinn, S. C. *et al.* Racial disparities in exposure, susceptibility, and access to health care in the US H1N1 influenza pandemic. *Am. J. Public Health* **101**, 285–293. https://doi.org/10.2105/AJPH.2009.188029 (2011).
26. Placzek, H. & Madoff, L. Effect of race/ethnicity and socioeconomic status on pandemic H1N1-related outcomes in Massachusetts. *Am. J. Public Health* **104**, e31-38. https://doi.org/10.2105/AJPH.2013.301626 (2014).
27. Virtual Pediatric Systems LLC. The VPS COVID-19 Dashboard: North American Pediatric ICUs. https://covid19.myvps.org/. Accessed 26 Sep 2020.
28. Centers for Disease Control and Prevention. COVID-NET: A Weekly Summary of US COVID-19 Hospitalization Data. https://gis.cdc.gov/grasp/COVIDNet/COVID19_5.html. Accessed 30 Oct 2020.
29. Kim, L. *et al.* Hospitalization rates and characteristics of children aged <18 years hospitalized with laboratory-confirmed COVID-19—COVID-NET, 14 States, March 1-July 25, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 1081–1088. https://doi.org/10.15585/mmwr.mm6932e3 (2020).
30. Hoang, A. *et al.* COVID-19 in 7780 pediatric patients: A systematic review. *EClinicalMedicine* **24**, 100433. https://doi.org/10.1016/j.eclinm.2020.100433 (2020).
31. Yonker, L. M. *et al.* Pediatric Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2): Clinical presentation, infectivity, and immune responses. *J. Pediatr.* **227**, 45–52. https://doi.org/10.1016/j.jpeds.2020.08.037 (2020).
32. Otto, W. R. *et al.* The epidemiology of severe acute respiratory syndrome coronavirus 2 in a pediatric healthcare network in the United States. *J. Pediatr. Infect. Dis. Soc.* **9**, 523–529. https://doi.org/10.1093/jpids/piaa074 (2020).
33. Poline, J. *et al.* Systematic severe acute respiratory syndrome coronavirus 2 screening at hospital admission in children: A French Prospective Multicenter Study. *Clin. Infect. Dis.* https://doi.org/10.1093/cid/ciaa1044 (2020).
34. Shekerdemian, L. S. *et al.* Characteristics and outcomes of children with coronavirus disease 2019 (COVID-19) infection admitted to US and Canadian pediatric intensive care units. *JAMA Pediatr.* **174**, 868–873. https://doi.org/10.1001/jamapediatrics.2020.1948 (2020).
35. Al Yazidi, L. S. *et al.* Epidemiology, characteristics and outcome of children hospitalized with COVID-19 in Oman: A multicenter cohort study. *Int. J. Infect. Dis.* **104**, 655–660. https://doi.org/10.1016/j.ijid.2021.01.036 (2021).
36. Zimmermann, P. & Curtis, N. Coronavirus infections in children including COVID-19: An overview of the epidemiology, clinical features, diagnosis, treatment and prevention options in children. *Pediatr. Infect. Dis. J.* **39**, 355–368. https://doi.org/10.1097/INF.0000000000002660 (2020).
37. Song, X. *et al.* Comparison of clinical features of COVID-19 vs seasonal influenza A and B in US children. *JAMA Netw. Open* **3**, e2020495. https://doi.org/10.1001/jamanetworkopen.2020.20495 (2020).
38. Gotzinger, F. *et al.* COVID-19 in children and adolescents in Europe: A multinational, multicentre cohort study. *Lancet Child. Adolesc. Health* **4**, 653–661. https://doi.org/10.1016/S2352-4642(20)30177-2 (2020).
39. Teherani, M. F. *et al.* Burden of illness in households with severe acute respiratory syndrome coronavirus 2-infected children. *J. Pediatr. Infect. Dis. Soc.* **9**, 613–616. https://doi.org/10.1093/jpids/piaa097 (2020).
40. Parcha, V. *et al.* Geographic variation in cardiovascular health among American adults. *Mayo Clin. Proc.* https://doi.org/10.1016/j.mayocp.2020.12.034 (2021).

41. Heald-Sargent, T. *et al.* Age-related differences in nasopharyngeal severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) levels in patients with mild to moderate coronavirus disease 2019 (COVID-19). *JAMA Pediatr.* https://doi.org/10.1001/jamapediatrics.2020.3651 (2020).

42. Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506. https://doi.org/10.1016/S0140-6736(20)30183-5 (2020).

43. Guan, W. J. *et al.* Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **382**, 1708–1720. https://doi.org/10.1056/NEJMoa2002032 (2020).

44. Grasselli, G. *et al.* Risk factors associated with mortality among patients with COVID-19 in intensive care units in Lombardy, Italy. *JAMA Intern. Med.* **180**, 1345–1355. https://doi.org/10.1001/jamainternmed.2020.3539 (2020).

45. Gundlapalli, A. V. *et al.* Death certificate-based ICD-10 diagnosis codes for COVID-19 mortality surveillance—United States, January–December 2020. *MMWR Morb. Mortal. Wkly. Rep.* https://doi.org/10.15585/mmwr.mm7014e2 (2021).

46. Ng, K. W. *et al.* Preexisting and de novo humoral immunity to SARS-CoV-2 in humans. *Science* **370**, 1339–1343. https://doi.org/10.1126/science.abe1107 (2020).

47. Bunyavanich, S., Do, A. & Vicencio, A. Nasal gene expression of angiotensin-converting enzyme 2 in children and adults. *JAMA* https://doi.org/10.1001/jama.2020.8707 (2020).

48. Xie, X., Chen, J., Wang, X., Zhang, F. & Liu, Y. Age- and gender-related difference of ACE2 expression in rat lung. *Life Sci.* **78**, 2166–2171. https://doi.org/10.1016/j.lfs.2005.09.038 (2006).

49. Mihalopoulos, M. *et al.* The resilient child: Sex hormones and COVID-19 incidence in pediatric patients. *J. Endocr. Soc.* https://doi.org/10.1210/jendso/bvaa106 (2020).

50. Wu, H. *et al.* Clinical and immune features of hospitalized pediatric patients with coronavirus disease 2019 (COVID-19) in Wuhan, China. *JAMA Netw. Open* **3**, e2010895. https://doi.org/10.1001/jamanetworkopen.2020.10895 (2020).

51. Krahn, G. L., Walker, D. K. & Correa-De-Araujo, R. Persons with disabilities as an unrecognized health disparity population. *Am. J. Public Health* **105**(Suppl 2), S198-206. https://doi.org/10.2105/AJPH.2014.302182 (2015).

52. Spong, C. Y. & Bianchi, D. W. Improving public health requires inclusion of underrepresented populations in research. *JAMA* **319**, 337–338. https://doi.org/10.1001/jama.2017.19138 (2018).

53. Turk, M. A., Landes, S. D., Formica, M. K. & Goss, K. D. Intellectual and developmental disability and COVID-19 case-fatality trends: TriNetX analysis. *Disabil. Health J.* **13**, 100942. https://doi.org/10.1016/j.dhjo.2020.100942 (2020).

54. Cummings, M. J. *et al.* Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: A prospective cohort study. *Lancet* **395**, 1763–1770. https://doi.org/10.1016/S0140-6736(20)31189-2 (2020).

55. Burke, R. M. *et al.* Symptom profiles of a convenience sample of patients with COVID-19—United States, January–April 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 904–908. https://doi.org/10.15585/mmwr.mm6928a2 (2020).

56. Kalra, R. *et al.* Implications of atrial fibrillation among patients with atherosclerotic cardiovascular disease undergoing noncardiac surgery. *Am. J. Cardiol.* **125**, 1836–1844. https://doi.org/10.1016/j.amjcard.2020.03.028 (2020).

57. Kalra, R., Patel, N., Doshi, R., Arora, G. & Arora, P. Evaluation of the incidence of new-onset atrial fibrillation after aortic valve replacement. *JAMA Intern. Med.* https://doi.org/10.1001/jamainternmed.2019.0205 (2019).

58. Harrison, S. L., Fazio-Eynullayeva, E., Lane, D. A., Underhill, P. & Lip, G. Y. H. Comorbidities associated with mortality in 31,461 adults with COVID-19 in the United States: A federated electronic medical record analysis. *PLoS Med.* **17**, e1003321. https://doi.org/10.1371/journal.pmed.1003321 (2020).

59. Singh, S. *et al.* Risk of severe coronavirus disease 2019 in patients with inflammatory bowel disease in the United States: A multicenter research network study. *Gastroenterology* https://doi.org/10.1053/j.gastro.2020.06.003 (2020).

60. Singh, S. & Khan, A. Clinical characteristics and outcomes of coronavirus disease 2019 among patients with preexisting liver disease in the United States: A multicenter research network study. *Gastroenterology* **159**, 768–771. https://doi.org/10.1053/j.gastro.2020.04.064 (2020).

61. Alkhouli, M., Nanjundappa, A., Annie, F., Bates, M. C. & Bhatt, D. L. Sex differences in case fatality rate of COVID-19: Insights from a multinational registry. *Mayo Clin. Proc.* **95**, 1613–1620. https://doi.org/10.1016/j.mayocp.2020.05.014 (2020).

62. Harrison, P. J., Luciano, S. & Colbourne, L. Rates of delirium associated with calcium channel blockers compared to diuretics, renin-angiotensin system agents and beta-blockers: An electronic health records network study. *J. Psychopharmacol.* **34**, 848–855. https://doi.org/10.1177/0269881120936501 (2020).

63. Taquet, M., Luciano, S., Geddes, J. R. & Harrison, P. J. Bidirectional associations between COVID-19 and psychiatric disorder: Retrospective cohort studies of 62 354 COVID-19 cases in the USA. *Lancet Psychiatry* https://doi.org/10.1016/S2215-0366(20)30462-4 (2020).

64. Hua-Gen Li, M., Hutchinson, A., Tacey, M. & Duke, G. Reliability of comorbidity scores derived from administrative data in the tertiary hospital intensive care setting: A cross-sectional study. *BMJ Health Care Inform.* https://doi.org/10.1136/bmjhci-2019-000016 (2019).

65. Walkey, A. J., Wiener, R. S., Ghobrial, J. M., Curtis, L. H. & Benjamin, E. J. Incident stroke and mortality associated with new-onset atrial fibrillation in patients hospitalized with severe sepsis. *JAMA* **306**, 2248–2254. https://doi.org/10.1001/jama.2011.1615 (2011).

66. Walkey, A. J. *et al.* Atrial fibrillation among Medicare beneficiaries hospitalized with sepsis: Incidence and risk factors. *Am. Heart J.* **165**, 949–955. https://doi.org/10.1016/j.ahj.2013.03.020 (2013).

67. Centers for Disease Control and Prevention. NCHS International Classification of Diseases, Tenth Revision, Clinical Modification (ICD-10-CM). https://www.cdc.gov/nchs/icd/icd10cm.htm. Accessed 25 Sep 2020.

68. Harris, C. R. *et al.* Array programming with NumPy. *Nature* **585**, 357–362. https://doi.org/10.1038/s41586-020-2649-2 (2020).

69. Haukoos, J. S. & Lewis, R. J. The propensity score. *JAMA* **314**, 1637–1638. https://doi.org/10.1001/jama.2015.13480 (2015).

70. Feudtner, C., Feinstein, J. A., Zhong, W., Hall, M. & Dai, D. Pediatric complex chronic conditions classification system version 2: Updated for ICD-10 and complex medical technology dependence and transplantation. *BMC Pediatr.* **14**, 199. https://doi.org/10.1186/1471-2431-14-199 (2014).

Acknowledgements

Author contributions

P.A. and V.P. contributed to the conceptualization, study design, data collection, carrying out the initial analyses, drafting the initial manuscript, and reviewing and revising the manuscript. K.S.B. and R.K. contributed to the drafting of the initial manuscript and critically reviewed the manuscript for important intellectual content. S.K. contributed to the data collection, data analyses, and critically reviewing the manuscript for important intellectual content. G.A. and L.B. contributed to the study design, critically reviewing the manuscript for important

intellectual content and revising the manuscript. All authors approved the final manuscript as submitted and agree to be accountable for all aspects of the work.

Competing interests

Dr. Kuranz is an employee of TriNetX, Inc, Cambridge, MA (USA). None of the other authors had any conflicts of interest or financial disclosures to declare.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/ 10.1038/s41598-021-89553-1.

Correspondence and requests for materials should be addressed to P.A.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.