Major article

Estimating pediatric cases of COVID-19 over time in the United States: Filling in a gap in public use data

Timothy L Wiemken PhD, MPH, FAPIC, FSHEA, CIC [a,*], Samson L Niemotka MS [b], Jacob A Clarke BS [c], Christopher G Prener PhD [d], Ana Santos Rutschman SJD [e]

[a] *Saint Louis University School of Medicine, Department of Internal Medicine, Division of Infectious Diseases, Allergy, and Immunology, Saint Louis, MO, USA*
[b] *Saint Louis University Institute for Vaccine Science and Policy, Data Science and Epidemiology Core, Saint Louis, MO, USA*
[c] *Saint Louis University School of Medicine, Saint Louis, MO, USA*
[d] *Saint Louis University School of Arts and Sciences, Department of Sociology and Anthropology, Saint Louis, MO, USA*
[e] *Saint Louis University School of Law, Center for Health Law Studies, Saint Louis, MO, USA*

Background: COVID-19 continues to disturb nearly all aspects of life, leaving us striving to reach herd immunity. Currently, only weekly standardized incidence rate data per age group are publicly available, limiting assessment of herd immunity. Here, we estimate the time-series case counts of COVID-19 among age groups currently ineligible for vaccination in the USA.
Methods: This was a secondary analysis of publicly available data. COVID-19 case counts by age groups were computed using incidence rate data from the CDC and population estimates from the US Census Bureau. We also created a web-based application to allow on demand analysis.
Results: A total of 78 weeks of data were incorporated in the analysis, suggesting the highest peak in cases within the 5−11-year age group on week ending 2021-01-09 (n = 61,095) followed by the 12-15-year age group (n = 58,093). As of July 24, 2021, case counts in the 5-11-year age group have expanded beyond other groups rapidly.
Discussion: This study suggests it is possible to estimate pediatric case counts of COVID-19. National agencies should report COVID-19 time series case counts for pediatric age cohorts. These data will enhance our ability to estimate the population at risk and tailor interventions accordingly.

INTRODUCTION

Coronavirus disease 2019 (COVID-19) continues to disturb nearly all aspects of life worldwide. Although prevention methods such as vaccination, face coverings, hand hygiene, and surface disinfection are effective, easy-to-institute, and readily available,[1-3] recent decreases in utilization of various interventions is concerning.[4,5] Despite this, a central approach to limit disease transmission, morbidity, and mortality is ensuring effective interventions are tailored to those at risk of infection and must continue to preserve public health.

With respect to COVID-19, we must continue to strive to meet herd immunity even if it is difficult to achieve.[6] Herd immunity results from when a population who mix or contact one another have sufficient immunity to prevent transmission of the agent causing the disease. This number has been described to be at least 60% for Severe Acute Respiratory Syndrome Coronavirus Type 2 (SARS-CoV-2), but the actual effective herd immunity threshold is difficult to compute. This difficulty may be explained by factors such as non-random mixing of the population, ignoring completely closed or nearly isolated populations, and an inability to determine both the longevity of immunity and overlap between immunity from natural infection and from vaccination. Because of these factors, we must rely on an assumption that the currently available and highly effective vaccines for COVID-19 provide enough long-term protection from SARS-CoV-2 infection to meet a herd immunity once at least 60% of the population is vaccinated.

* Address correspondence to Timothy L Wiemken, PhD, MPH, FAPIC, FSHEA, CIC, Saint Louis University School of Medicine, Department of Internal Medicine, Division of Infectious Diseases, Allergy, and Immunology, Data Science and Epidemiology, Institute for Vaccine Science and Policy, Infectious Diseases Epidemiology, SSM Saint Louis University Hospital, Systems Infection Prevention Center, 1100 South Grand Blvd #817, Doisy Research Center, St. Louis, MO 63104.
E-mail address: timothy.wiemken@health.slu.edu (T.L. Wiemken).
Conflicts of interest: None to report.

Currently, three vaccines are approved, or authorized for emergency use for prevention of COVID-19 in the United States for various age groups: Pfizer/BioNTech (≥12 years of age), Moderna (≥18 years of age), and Johnson & Johnson/Janssen (≥18 years of age), with Pfizer BioNTech being the only fully approved by the United States Food and Drug Administration (FDA). To reach herd immunity, approval or emergency use authorization must be provided to younger age groups quickly, since approximately 15% (≤12 years of age) to 20% (≤15 years of age) of the United States population is ineligible for vaccination from all or some of the vaccines available for adults.

To begin understanding the status of herd immunity, we must understand immunity due to SARS-CoV-2 and know the population who is immune and at risk. To determine the population at risk, we need to understand the trajectory of cases, particularly across groups ineligible for vaccine. Currently, only weekly standardized incidence rate data per age group are publicly available,[7] limiting community assessment of herd immunity.

The objective of this study was to utilize publicly available data to estimate the time-series case counts of COVID-19 among age groups currently ineligible for vaccination in the United States.

METHODS

This was a secondary analysis of publicly available data from week ending March 07, 2020 through August 21, 2021. To compute case counts over time and by age group, we first obtained COVID-19 incidence rate data per 100,000 population by age group from the Centers for Disease Control and Prevention (CDC).[7] These data are published pre-standardized and therefore do not allow for computation of total caseload directly. We then obtained calendar year 2020 and 2021 single age population estimates from the United States Census Bureau.[8] Using these two data elements, we were able to back compute the de-standardized rate using the following formula:

*(Age-specific COVID-19 rate per 100,000 population for 2020 or 2021 * Age-specific estimated population for the same year) / 100,000.*

This formula provided an estimate of the case count by age group using the population estimates specific for the year for which the incidence rates were published.

Next, we obtained overall (not time-series) COVID-19 case counts by age-group from the CDC[9] and computed a percent over or underestimate from our estimates to the actual reported data by age group. These percentages were used to correct the weekly case counts by adding or subtracting the specific percentage over/underestimate for each time-period and age group.

We then computed percent changes from the prior week for each count by age group using the formula:

*((current case count − prior case count) / current case count) *100*

Next, we computed herd immunity estimates for the age groups currently ineligible for COVID-19 vaccination in the United States (0-4 and 5-11 age groups), simulating different scenarios of under-reporting.

Finally, we created a web-based application to allow interested readers to visualize data.

RESULTS

A total of 78 weeks of data were incorporated in the analysis, including week ending 2020-03-07 through 2021-08-28. Total estimated case counts were close to actual case counts across the three age groups and were closer to published data after correction for over/under estimation except for the 12-15-year age group (Table 1). Overall reported case counts were highest among the 5-11 age group (n = 1,362,512) followed by the 12-15-year age group (n = 1,182,692), and finally the 0-4-year age group (n = 693,100). Time series plots suggest the highest peak in cases within the 5-11-year age group on week ending January 09, 2021 (n = 61,095) followed closely by the 12-15-year age group (n = 58,093); though incidence rate during this time was higher in the 12-15-year age group. As of approximately July 24, 2021, case counts in the 5-11-year age group have expanded beyond 0-4 and 12-15 year age groups more rapidly than at any other time period during the pandemic (Fig 1).

Percent change week-on-week varied substantially with a mean percent change of 2.3% over the time series and has experience more week-to-week variability since April 2021 (Fig 2).

Herd immunity estimates for vaccine-ineligible age groups can be found in Table 2. Even with 100% under-reporting the level of immunity due to natural infection (in the best-case scenario of perfect, non-waning immunity from natural infection), only 4.2% of these children are currently immune. To bring these groups to an 80% immunity threshold, a value more realistic than prior calculations given the increased transmissibility of the delta variant,[10] there would need to be a 1900% under-reporting bias.

The web-based tool for on-demand visualization can be found at: https://surveillance.shinyapps.io/covid_agegroups. This tool allows for rapid visualization of all age groups with data available for estimated case counts, corrected estimated case counts, and percent differences.

DISCUSSION

This study suggests it is possible to estimate with a reasonable degree of accuracy, the case counts of COVID-19 in younger age groups. Given national data repositories do not report pediatric COVID-19 cases by age group, the practitioner's ability to discern risks and target intervention to reduce transmission is limited. Beyond our estimates, in an impossible scenario of perfect non-waning immunity from infection, only 4.2% of the 0-11 aged population is immune. Even if this estimate is 100% underestimated, it leaves a massive at-risk population in critical need of vaccine.

Interestingly, time-series evaluation of case counts modeled each other across the three age groups closely, suggesting proportional and equivalent transmission and infection across age groups given the population at risk. Percent change estimates were also nearly identical week-on-week for all three age groups further adding credence to similar exposure, risk, and infection profiles across age groups.

This study does not suggest that incidence rate data are not useful. In fact, arguably for most surveillance, incidence rates are far more useful as they account for the population at risk. A close evaluation of the incidence rate data by age groups currently ineligible for vaccination underscores this point.[7] Throughout the pandemic, the 12-15-year old age group had higher incidence rates than did both 0-4 and 5-11 age groups, despite having lower total cases than the 5-11 cohort. However, after Pfizer BioNTech vaccine was authorized for emergency use in the 12-15-year old age group (May 10, 2021), rates

Table 1
Week-on-week estimated case counts and corrected case counts with totals by age group, United States, weeks ending March 07, 2020 through August 28, 2021

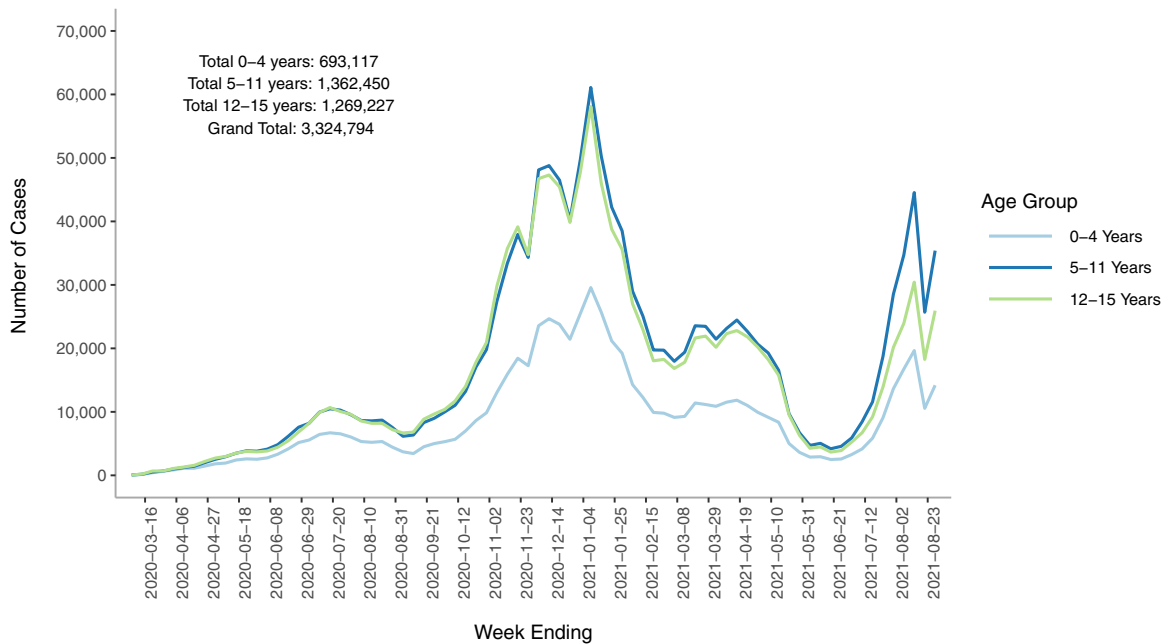| Age Group	Total Reported COVID-19 Case Count[9]	Estimated Case Count (Percent Difference from Actual)	Corrected Estimated Case Count (Percent Difference from Actual)
0-4	639,100	710,307 (10.0%)	693,117 (7.8%)
5-11	1,362,512	1,319,053 (-3.3%)	1,362,450 (<0.00%)
12-15	1,182,692	1,160,171 (-1.9%)	1,269,227 (6.8%)

Fig 1. Time series plot of estimated COVID-19 case counts (corrected) by age group, United States, week ending March 07, 2020 through August 28, 2021.

of disease between ages 5-11 and 12-15 have been much closer with approximately 3.5% of the 12-15 year old age group currently fully vaccinated.[11]

This study has limitations. Since COVID-19 surveillance is largely passive in the United States, it is possible that underreporting is common.[12] However, recent seroprevalence studies among younger age groups suggest low infection rates[13,14] resulting in an unknown impact of this factor on our estimates. Although we were able to correct our estimates to align with the total reported COVID-19 cases more closely, they remain variable, and our corrections may not be directly applicable to each weekly report. This could bias the visualizations and/or the total counts of those immune.

In conclusion, national agencies should consider reporting COVID-19 time series case counts for pediatric age cohorts along with the standardized rates. These data will enhance public health practitioners' abilities to estimate the population at risk and/or immune and tailor interventions accordingly.
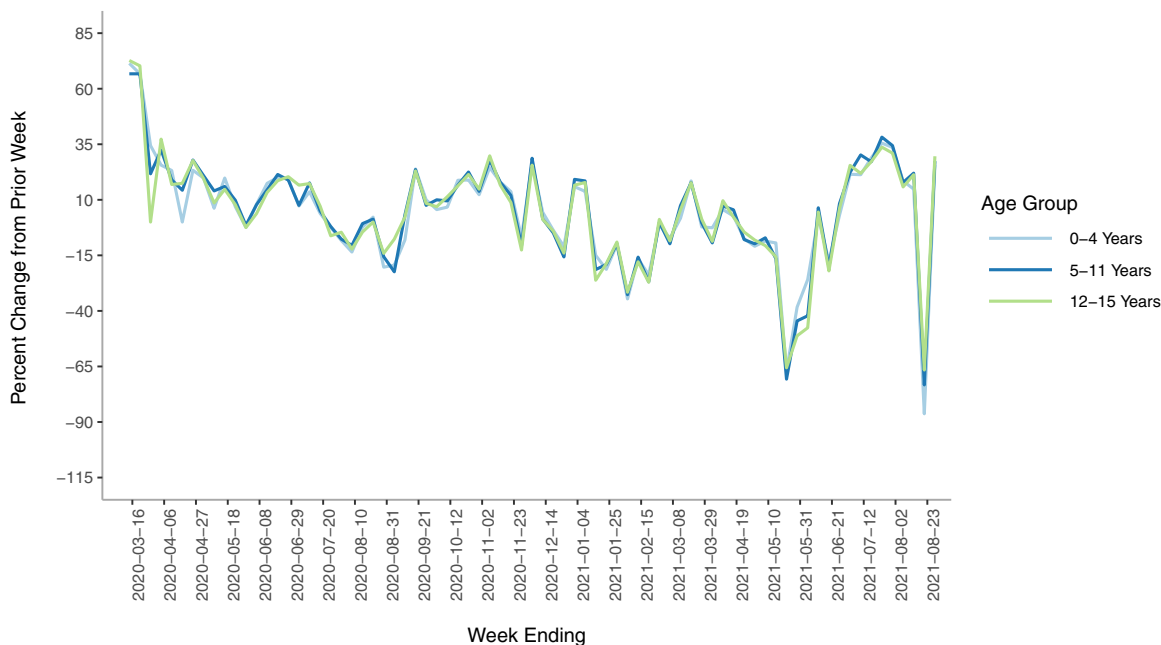

Fig 2. Time series plot of estimated percent difference week-on-week COVID-19 case counts (corrected) by age group, United States, week ending March 07, 2020 through August 28, 2021.

Table 2
Herd immunity estimates for 0-4- and 5-11-year-old, COVID-19 vaccine ineligible age cohorts through week ending August 28, 2021

Age Group	Projected 2021 Population	Total Reported COVID-19 Case Count (Percent of Population Immune)	Percent Under Reporting					
			0%	10%	25%	50%	75%	100%
			Estimated Case Count (Percent of Population Immune)	Estimated Case Count (Percent of Population Immune)	Estimated Case Count (Percent of Population Immune)	Estimated Case Count (Percent of Population Immune)	Estimated Case Count (Percent of Population Immune)	Estimated Case Count (Percent of Population Immune)
0-4	20,601,023	639,100 (3.1)	671,055 (3.3)	7030,10 (3.4)	798,875 (3.9)	958,650 (4.7)	1,118,425 (5.4)	1,278,200 (6.2)
5-11	28,368,818	1,362,512 (4.8)	1,430,638 (5.0)	1,498,763 (5.3)	1,703,140 (6.0)	2043,768 (7.2)	2,384,396 (8.4)	2,725,024 (9.6)
Total	48,969,841	2,055,612 (4.2)	2,158,397 (4.4)	2,261,173 (4.6)	2,569,515 (5.3)	3,083,418 (6.3)	3,597,321 (7.4)	4,111,224 (8.4)

References

1. Baker RE, Park SW, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. The impact of COVID-19 nonpharmaceutical interventions on the future dynamics of endemic infections. *Proc Natl Acad Sci U S A*. 2020;117:30547–30553.
2. Wiemken T. Fomite transmission among respiratory viruses and the importance of low-level surface disinfection. Harv Public Health Rev. 202;28.
3. Pradhan D, Biswasroy P, Kumar Naik P, Ghosh G, Rath G. A review of current interventions for COVID-19 prevention. *Arch Med Res*. 2020;51:363–374.
4. Crane MA, Shermock KM, Omer SB, Romley JA. Change in reported adherence to nonpharmaceutical interventions during the COVID-19 pandemic, April-November 2020. *JAMA*. 2021;325:883–885.
5. USA Today. COVID-19 restrictions: Map of COVID-19 case trends, restrictions and mobility.
6. Kwok KO, McNeil EB, Tsoi MTF, Wei VWI, Wong SYS, Tang JWT. Will achieving herd immunity be a road to success to end the COVID-19 pandemic? *J Infect*. 2021;83:381–412.
7. Centers for Disease Control and Prevention. COVID Data Tracker: COVID-19 Weekly Cases and Deaths per 100,000 Population by Age, Race/Ethnicity, and Sex 2021 [updated August 21, 2021]. Accessed November 18, 2021. https://covid.cdc.gov/covid-data-tracker/#demographicsovertime.
8. United States Census Bureau. 2017 National Population Projections Datasets 2021 [updated February 20, 2020]. Accessed November 18, 2021. https://www.census.gov/data/datasets/2017/demo/popproj/2017-popproj.html.
9. Centers for Disease Control and Prevention. COVID Data Tracker: Demographic Trends of COVID-19 cases and deaths in the US reported to CDC 2021 [updated August 30, 2021]. Accessed November 18, 2021. https://covid.cdc.gov/covid-data-tracker/#demographics.
10. Campbell F, Archer B, Laurenson-Schafer H, et al. Increased transmissibility and global spread of SARS-CoV-2 variants of concern as at June 2021. *Euro Surveill*. 2021;26:2100509.
11. Centers for Disease Control and Prevention. COVID Data Tracker: Demographic Characteristics of People Receiving COVID-19 Vaccinations in the United States 2021 [updated August 31, 2021]. Accessed November 18, 2021. https://covid.cdc.gov/covid-data-tracker/#vaccination-demographic.
12. Wu SL, Mertens AN, Crider YS, et al. Substantial underestimation of SARS-CoV-2 infection in the United States. *Nat Commun*. 2020;11:4507.
13. Smith BK, Janowski AB, Danis JE, et al. Seroprevalence of SARS-CoV-2 antibodies in children and adults in St. Louis, Missouri, USA. *mSphere*. 2021;6:e01207–20.
14. Sola AM, David AP, Rosbe KW, Baba A, Ramirez-Avila L, Chan DK. Prevalence of SARS-CoV-2 infection in children without symptoms of coronavirus disease 2019. *JAMA Pediatr*. 2021;175:198–201.