

Estimated incidence of COVID-19 illness and hospitalization — United States, February–September, 2020

Heather Reese^{1,2}, A. Danielle Iuliano^{1,3}, Neha N. Patel¹, Shikha Garg^{1,3}, Lindsay Kim^{1,3}, Benjamin J. Silk^{1,3}, Aron J. Hall¹, Alicia Fry^{1,3}, Carrie Reed¹

¹ COVID-19 Emergency Response, Centers for Disease Control and Prevention

² Epidemic Intelligence Service, Centers for Disease Control and Prevention

³ U.S. Public Health Service

Corresponding Author:

Carrie Reed

Centers for Disease Control and Prevention

1600 Clifton Rd NE, MS H24-7

Atlanta, GA 30329

Email: creed1@cdc.gov

Summary: We estimated the cumulative incidence of COVID-19 in the U.S. population by adjusting confirmed case counts for sources of under-detection. Through September 30, 2020, an estimated 2.4 million hospitalizations, 44.8 million symptomatic illnesses, and 52.9 million SARS-CoV-2 infections may have occurred.

Accepted Manuscript

Abstract

Background: In the United States, laboratory confirmed coronavirus disease 2019 (COVID-19) is nationally notifiable. However, reported case counts are recognized to be less than the true number of cases because detection and reporting are incomplete and can vary by disease severity, geography, and over time.

Methods: To estimate the cumulative incidence SARS-CoV-2 infections, symptomatic illnesses, and hospitalizations, we adapted a simple probabilistic multiplier model. Laboratory-confirmed case counts that were reported nationally were adjusted for sources of under-detection based on testing practices in inpatient and outpatient settings and assay sensitivity.

Results: We estimated that through the end of September, 1 of every 2.5 (95% Uncertainty Interval (UI): 2.0–3.1) hospitalized infections and 1 of every 7.1 (95% UI: 5.8–9.0) non-hospitalized illnesses may have been nationally reported. Applying these multipliers to reported SARS-CoV-2 cases along with data on the prevalence of asymptomatic infection from published systematic reviews, we estimate that 2.4 million hospitalizations, 44.8 million symptomatic illnesses, and 52.9 million total infections may have occurred in the U.S. population from February 27–September 30, 2020.

Conclusions: These preliminary estimates help demonstrate the societal and healthcare burdens of the COVID-19 pandemic and can help inform resource allocation and mitigation planning.

Background

In the United States, the earliest known patients with coronavirus disease 2019 (COVID-19), the disease caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection, were associated with travel to affected countries or known contact with other infected persons [1]. By February 2020, persons with SARS-CoV-2 infection in the U.S. and no known exposure were detected [2]. Between February 27–September 30, 2020, nearly 6.9 million laboratory-confirmed cases of domestically acquired infections were detected and reported nationally.

Persons with laboratory-confirmed SARS-CoV-2 infection reported through national surveillance do not represent all infected persons in the U.S. Seroprevalence studies have shown a higher level of SARS-CoV-2 infection than has been reflected by confirmed case counts [2-7]. Most unreported infections were asymptomatic or mildly ill people who recovered without seeking medical care or testing [8-10]. However, even persons with SARS-CoV-2 infection in medical settings may not be tested or nationally reported as confirmed cases. Limited availability of tests, reagents, and laboratory capacity reduced case detection, in addition patients may have avoided medical care settings or presented with non-specific symptoms and not been suspected to have SARS-CoV-2 infection. Furthermore, not all infected persons will test positive because of assay sensitivity, timing of specimen collection, or specimen quality [11]. Factors involved in detecting and reporting cases may vary by age, geographically, over time, across healthcare settings, and by severity of disease. Finally, some people may be infected with SARS-CoV-2 and never show clinical symptoms; these asymptomatic persons would be even less likely to be detected [9, 10].

To better estimate the U.S. incidence of SARS-CoV-2 infection since the beginning of the pandemic, we adapted a probabilistic multiplier model to adjust nationally reported counts of confirmed cases for various sources of

under-detection [12]; this model estimates total SARS-CoV-2 infections, symptomatic illnesses, and hospitalized patients in the U.S. population from February 27, 2020–September 30, 2020.

Methods

Reported confirmed cases

Persons with laboratory confirmed SARS-CoV-2 infection by molecular diagnostics are reported to CDC through the Nationally Notifiable Disease Surveillance System (NNDSS) at the person level or as aggregate counts at the reporting jurisdiction level (e.g., state, territory, New York City, District of Columbia) [13, 14]. The NNDSS uses a standardized case report form, including state of residence, age, hospitalization admission, and other demographic and clinical characteristics. Given data entry delays and incomplete national reporting, jurisdictions reported aggregated counts daily for the previous day. Probable, asymptomatic, and travel-associated cases were excluded from counts of confirmed cases used in this analysis.

Analytic methods

We applied a probabilistic multiplier model to adjust the reported numbers of confirmed symptomatic cases for factors affecting detection of persons with SARS-CoV-2 infection, a method previously used to estimate the incidence of H1N1pdm09 during the 2009 influenza pandemic [12]. This method uses confirmed cases and data on case detection and the asymptomatic fraction to estimate the cumulative number of hospitalized patients

with SARS-CoV-2 infection, the total number with symptomatic illness, and the total number of infected persons (Figure 1).

To account for variability in detection of SARS-CoV-2 we stratified reported cases into hospitalized and non-hospitalized symptomatic cases, and further by age group (0–4 years, 5–17 years, 18–49 years, 50–64 years, 65 years and older), time period when the case was reported (February–March, April–May, June–July, August–September), and U.S. Department of Health and Human Service (HHS) region[15]. Age group was imputed for cases with missing birth date according to the age distribution within each HHS region and reporting time period. If hospitalization status was missing, we imputed the percentage of patients who were hospitalized based on reported cases with complete data within by age group, HHS region, and reporting time period. More details on this process are available in the supplemental material.

We adjusted case counts for three factors that affected national case detection of symptomatic cases: if a patient is symptomatic, they may not have sought medical attention or testing for their illness (parameter C); if a patient sought medical care, they may not have had a SARS-CoV-2 test completed (parameter B); or if a patient was tested, the SARS-CoV-2 assay used may result in a false negative result due to its sensitivity to detect SARS-CoV-2 in the specimen (parameter A). We used several data sources to describe these factors (Table 2), with under-detection multipliers calculated as an inverse of the product of factors A-C. Each multiplier was calculated within strata of hospitalization status, age group, reporting time period, as data were available, and applied to the relevant stratified cases counts to estimate a number of symptomatic cases within that strata.

After adjustment, we summed the strata to a number of estimated symptomatic cases and applied one more source of under-detection – a person infected with SARS-CoV-2 may never show clinical symptoms (parameter D) – to estimate the number of total infections in the population.

For all parameters and strata, we included a range of values; estimates were calculated using Latin hypercube sampling with 10,000 iterations, with 95% uncertainty intervals estimated as the 2.5th and 97.5th percentile range. Population rates were estimated using bridged-race population estimates from CDC Wonder [16]. Analyses were completed in R (version 3.6.1).

Sources of under-detection of cases

Parameter A. SARS-CoV-2 assay sensitivity

Patients infected with SARS-CoV-2 may not always test positive. Sensitivity of approved molecular diagnostic assays may be affected by the limits of detection of specific assays [10], specimen quality, source, handling, and timing of collection [11]. In a systematic review, 2%–21% of patients ultimately confirmed to have SARS-CoV-2 infection did not have a positive result unless multiple tests were performed over several days [17]. This review was used to estimate the probability that a specimen with SARS-CoV-2 will test positive (Table 2). For simplicity, since reported assay specificity has been high with false positive results ranging between 1-4% [18, 19], we did not adjust for potential false positives.

Parameter B. SARS-CoV-2 assay ordered and test completed

Patients with SARS-CoV-2 infection who are not tested with molecular assays are not included in confirmed case counts. To characterize testing probabilities, we used data from two sources on healthcare visits and SARS-CoV-

2 testing, and estimated this parameter separately for hospitalized and non-hospitalized patients. To capture the variability in testing practices across data sources, we represented this parameter using a beta PERT distribution centered on the median value and ranging between the minimum and maximum values reported across both data sources within each stratum of age (Table 2). The beta PERT distribution is a continuous probability distribution, which emphasizes the most likely values in an acceptable range of parameter values (i.e., more often drawing closer to middle value of the interval with a smaller probability on the extremes of the interval).

The first source of data was the IBM Watson Health Explorys electronic health record (EHR) database (IBM, Armonk, NY), which includes >39 health system partners across the country. We identified visits with an ICD-10 diagnosis or SNOMED code that indicated an acute respiratory illness (ARI) (Supplemental Table 5) and the number of those with evidence of SARS-CoV-2 test results from LOINC codes for SARS-CoV-2 RT-PCR tests (Supplemental Table 6). For each setting (inpatient, outpatient ED), visits and tests performed were aggregated into strata for time period and age group.

We also included rates of testing in the COVID Near You (CNY) survey platform. CNY is a website application where participants can self-report symptoms, healthcare seeking behaviors, and SARS-CoV-2 testing information [20-22]. COVID-like illness (CLI) was defined using self-reported presence of shortness of breath or cough, or two or more of: self-reported fever, chills, sore throat, body ache, headache, or loss of taste or smell. Proportions of individuals who self-reported receiving a SARS-CoV-2 test among those who sought care for CLI were estimated for each time period with available data by HHS region, and age group (Table 2, Supplemental Table 4).

Parameter C. Symptomatic patient seeks care/testing

A symptomatic person with SARS-CoV-2 infection will not be included in confirmed case counts if they never sought medical attention or testing services. To estimate healthcare seeking, we used data obtained from both

CNY and Flu Near You (FNY) [23], which has conducted participatory surveillance for influenza-like illnesses since 2011, to better capture the full time period and differences between participants of the two systems. We considered a range of symptomatic illness including: (1) CLI as described above, but excluding loss of taste or smell for FNY, which was not captured in that platform; (2) a more specific case definition of fever, and either cough or shortness of breath; and, (3) a broader case definition of at least one of fever, cough, or shortness of breath. Among patients who met the given case definition, we calculated the proportion that reported visiting a doctor's office, urgent care clinic, outpatient clinic, emergency department, testing center, telemedicine, or other healthcare setting for symptoms. Care seeking proportions were included using a beta PERT distribution of the median and range of values across the three case definitions and two data sources, stratified by report date and age group (Table 2, Supplemental Table 2).

Parameter D. Patient is symptomatic if infected with SARS-CoV-2

Some people infected with SARS-CoV-2 do not experience symptoms [24]. To estimate the number of infections in the population, we adjusted the sum of hospitalized and symptomatic non-hospitalized cases based on the the proportion of persons with confirmed COVID-19 and no symptoms from a meta-analysis of available literature (Table 2) [17].

Results

National case reporting

During February 27–September 30, 2020, there were 6,891,764 confirmed cases of symptomatic COVID-19 acquired domestically and reported nationally through individual or aggregate case counts. We estimated that approximately 14% of these patients had been hospitalized, with variation by age group, case report date, and HHS region (Table 1).

Hospitalized cases

We estimated 2.5 (95% uncertainty interval [UI]: 2.0–3.1) SARS-CoV-2 hospitalizations in the population for each hospitalized case reported nationally, with variations by age group, HHS region, and report date. Under-detection multipliers decreased over time and were consistently highest among children (Supplemental Table 3).

Adjusting case counts by HHS region, age group, and report date, we estimated a total of 2,397,777 (95% UI: 2,053,156–2,855,843) hospitalizations with SARS-CoV-2 infection (Table 3), or 733 hospitalizations per 100,000 population. The highest rates of hospitalization were among patients aged ≥ 65 years (1,950/100,000), lowest among children 5–17 years of age (83/100,000). Estimates varied geographically: 236/100,000 in HHS Region 10 to 2,440/100,000 in HHS Region 2.

Non-hospitalized symptomatic illnesses

We estimated 7.1 (95% UI: 5.8–9.0) non-hospitalized symptomatic illnesses for every one non-hospitalized case reported nationally, with variation by age group, HHS region, and report date. Under-detection multipliers decreased over time and were consistently highest among children (Supplemental Table 3).

We summed the estimated hospitalized (Table 3) and non-hospitalized (Supplemental Table 5) illness for a total of 44.8 million symptomatic illnesses (Table 4). The highest rates of symptomatic illness were among adults 18–49 years old (18,162/100,000); lowest among children aged 0–4 years (5,777/100,000). Estimates varied geographically: 8,282/100,000 in HHS Region 10 to 26,705/100,000 in HHS Region 2.

Total infections

Using age-stratified estimates of the proportion of infections that remain asymptomatic, we estimated that the nationally reported cases during February–September may represent a total of 52,885,526 (95% UI: 42,527,569–66,810,205) SARS-CoV-2 infections in the U.S. population, with highest infection rates among persons aged 18–49 years (Table 5). This indicates that 1 in 7.7, or 13% of total infections were identified and reported. Detection varied by age, with lower detection rates among children, but with improvements over time (Supplemental Table 4).

Conclusions

We estimated that nearly 53 million SARS-CoV-2 infections, including 42 million symptomatic illnesses and 2.4 million associated hospitalizations, may have occurred in the U.S. through September 30, 2020; with variation by geographic region, age group, and time. These preliminary estimates demonstrate the large incidence of disease in the U.S. population and better quantify the impact of the COVID-19 pandemic on the healthcare system and society, and will be updated as more data on under-detection become available.

From past experiences with influenza [25], another respiratory virus associated with a large proportion of mild illness and an overlapping clinical syndrome with COVID-19, laboratory-confirmed cases reported through surveillance systems underestimate total infections. We adapted our current approach from methods to estimate the influenza A/H1N1pdm09 prevalence in the U.S. during the 2009 pandemic[12]. Our preliminary estimates indicate approximately 1 in 8, or 13%, of total SARS-CoV-2 infections were recognized and reported through the end of September. Similarly, a recent serologic survey of SARS-CoV-2 antibodies in 10 geographically diverse U.S. sites from March 23–May 12 of 2020 estimated that the total number of SARS-CoV-2 infections was

at least 10 (range by U.S. site: 6–24) for every reported case [3], with improvements in this ratio by later time points. Severe cases were more likely to be detected and reported; we estimated 2.5 hospitalized patients for each hospitalized case reported. In the Explorys EHR data, the proportion of ICU patients tested for SARS-CoV-2 was >90% by the end of September, though testing remained lower among other inpatients with ARI, and even lower for ARI visits in outpatient settings (Supplemental Figure 1).

For comparison, COVID-NET is an active, population-based surveillance system for laboratory-confirmed SARS-CoV-2-associated hospitalizations in defined areas of 14 states [26]. While direct comparisons with COVID-NET are imperfect due to the narrower geographic area of the surveillance sites, in 10 of the 14 sites, our estimated hospitalization rates by region were 1.5–3.5 times higher than the reported rates from individual sites within those regions by the end of September, similar to the range of our estimated under-detection multiplier for confirmed hospitalizations. Likewise, COVID-NET showed similar trends across age; adults aged ≥ 65 years had 5–6 times higher rates of hospitalizations than younger adults aged 18–49 years [27]. Both also showed lower hospitalization rates among children [28, 29].

For comparison of population-level incidence of infection, the estimated 36 million infections represent approximately 16% of the U.S. population, ranging from 9%–31% across regions of the country. This is higher than seroprevalence estimates from a nationwide commercial laboratory seroprevalence survey, which found that 1%–22% of various state populations had antibodies to SARS-CoV-2 by early August, though our estimates include two more months of circulation [31]. There remain uncertainties in the interpretation of seroprevalence estimates, including how they vary by the population surveyed, the serologic assays used, the proportion of infected cases with a detectable antibody response, and how long antibody detection persists after infection. Additional studies and sources of data on population-based incidence will help resolve these concerns and provide better national estimates of illness and infection.

We recognize that our model has limitations. From almost a decade of monitoring data on testing practices for influenza [32, 33], testing rates and the use of more sensitive molecular testing has varied by jurisdictions, care settings, age, and disease severity [34]. The availability and use of testing for SARS-CoV-2 has changed rapidly over time; thus far, data on the proportion of persons who are tested for COVID-19 and how this varies across all the previously described factors remains limited. Although data on testing by time, healthcare setting, and age was available, it lacked the coverage to allow for geographic-specific model inputs. These data limitations could have resulted in overestimation of cases from areas with higher testing rates, including some hospitals that are performing universal testing, or have more outpatient testing facilities and active contact tracing. Likewise, we may underestimate in areas with lower testing and contact tracing. Additionally, some infections, such as those among healthcare workers or from outbreaks in congregate residential settings, may be more likely to be tested and nationally reported compared with the general population, and could overestimate non-hospitalized cases and infections. We continue to seek information on the proportion of cases and testing rates in various settings to improve estimates. With limited but growing information regarding the spectrum of clinical manifestations from SARS-CoV-2 infection, there could be a lower index of suspicion of COVID-19 for patients who present with nonspecific and non-respiratory symptoms; these cases may be less likely to be detected and reported. All of this highlights the importance of having data to monitor the proportions of patients with different clinical syndromes who are being tested for SARS-CoV-2 infection in a variety of healthcare and geographic settings, and not just total numbers of tests performed. Finally, in some heavily affected areas, the size of the outbreaks exceeded capacities to complete detailed case reporting, including patient age and hospitalization status. For cases with missing hospitalization status, we imputed the proportion of reported cases that were hospitalized from the subset with complete data, but it is unclear if age and hospitalization status were missing at random[35]. If not random, and the data were more complete for hospitalized patients, the true hospitalization ratio would be lower than we imputed, and the number of hospitalized cases would be lower than we estimated. Furthermore, this was hospitalization status at the time of the case report, and would miss those

diagnosed as an outpatient but became hospitalized after they were reported as a case; thus our estimates of hospitalization may be an underestimate.

Despite these limitations, our model provides a relatively simple approach to illustrate why there are more persons who have had a SARS-CoV-2 infection than the reported confirmed case counts at multiple levels of disease severity. We used data currently available to provide a preliminary estimate of the overall incidence of SARS-CoV-2 infection, illness, and hospitalization in the U.S. CDC is actively working on refining methods to synthesize information across multiple data sources to better describe the national burden of SARS-CoV-2 infection on an ongoing basis and will update estimates as data become available.

In summary, we estimated that in the U.S. through September 30, 2020 there were approximately 53 million total SARS-CoV-2 infections, including 42 million symptomatic illnesses and 2.4 million hospitalizations, with large variations by age group and geographic area. This indicates that approximately 84% of the U.S. population has not yet been infected and thus most of the country remains at risk, despite already high rates of hospitalization. Improved estimates of SARS-CoV-2 infections, symptomatic illnesses, and hospitalizations over time, are critical to our understanding of the severity and burden of this new virus..

Accepted Manuscript

NOTES

Acknowledgements:

We would like to acknowledge the following additional individuals for their contribution to this manuscript. We thank Ryan Threlkel and Catherine Bozio at CDC for assistance on various parts of the analysis. We appreciate the efforts of John Brownstein, Jared Hawkins, Benjamin Rader, Kara Sewalk, Autumn Gertz, Christopher Rempel, Christina Astley, Thomas McMennamin, and Lakshmi Yajurvedi from Boston Children's Hospital for their work with COVID Near You and Flu Near You. We would like to thank Teresa Gibson, Bill Saunders, Brian Goodness, Kay Miller, Sarah Bloemers, Tim Burrell, and Shannon Harrer from IBM Watson for their assistance with the Explorys Health Record Database.

Disclaimer: The findings and conclusions in this report are those of the author(s) and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Accepted Manuscript

Accepted Manuscript

References

1. Bajema, K.L., et al., *Persons Evaluated for 2019 Novel Coronavirus - United States, January 2020*. MMWR Morb Mortal Wkly Rep, 2020. **69**(6): p. 166-170.
2. Team, C.C.-R., et al., *Evidence for Limited Early Spread of COVID-19 Within the United States, January-February 2020*. MMWR Morb Mortal Wkly Rep, 2020. **69**(22): p. 680-684.
3. Havers, F.P., et al., *Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020*. JAMA Intern Med, 2020.
4. Zhao, J., et al., *Antibody responses to SARS-CoV-2 in patients of novel coronavirus disease 2019*. Clin Infect Dis, 2020.
5. CDC. *Large-scale Geographic Seroprevalence Surveys*. 2020; Available from: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/geographic-seroprevalence-surveys.html>.
6. Bendavid, E., et al., *COVID-19 Antibody Seroprevalence in Santa Clara County, California*. medRxiv, 2020.
7. Sood, N., et al., *Seroprevalence of SARS-CoV-2-Specific Antibodies Among Adults in Los Angeles County, California, on April 10-11, 2020*. JAMA, 2020.
8. Davies, N.G., et al., *Age-dependent effects in the transmission and control of COVID-19 epidemics*. Nat Med, 2020.
9. Dong, Y., et al., *Epidemiology of COVID-19 Among Children in China*. Pediatrics, 2020. **145**(6).
10. Mizumoto, K., et al., *Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020*. Euro Surveill, 2020. **25**(10).
11. Woloshin, S., N. Patel, and A.S. Kesselheim, *False Negative Tests for SARS-CoV-2 Infection - Challenges and Implications*. N Engl J Med, 2020.
12. Reed, C., et al., *Estimates of the prevalence of pandemic (H1N1) 2009, United States, April-July 2009*. Emerg Infect Dis, 2009. **15**(12): p. 2004-7.
13. CDC. *2020 Interim Case Definition, Approved April 5, 2020*. 2020; Available from: <https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/>.
14. CDC. *2020 Interim Case Definition, Approved August 5, 2020*. 2020; Available from: <https://wwwn.cdc.gov/nndss/conditions/coronavirus-disease-2019-covid-19/case-definition/2020/08/05/>.
15. DHHS. *HHS Regional Offices*. 2014; Available from: <https://www.hhs.gov/about/agencies/iea/regional-offices/index.html>.

16. CDC. *CDC Wonder: Bridged-Race Population Estimates*. 2020; Available from: <https://wonder.cdc.gov/bridged-race-population.html>.
17. Arevalo-Rodriguez, I., et al., *False-negative results of initial RT-PCR assays for COVID-19: A systematic review*. medRxiv, 2020.
18. Cohen, A.N. and B. Kessel, *False positives in reverse transcription PCR testing for SARS-CoV-2*. medRxiv, 2020.
19. Mayers, C. and K. Baker, *Impact of false-positives and false-negatives in the UK's COVID-19 RT-PCR testing programme*. 2020, Scientific Advisory Group for Emergencies.
20. *COVID Near You*. 2020; Available from: <https://covidnearyou.org/us/en-US/>
https://figshare.com/articles/online%20resource/COVIDNearYou_COVID-19_Syndromic_Surveillance_Platform/12732857/1.
21. Chan, A.T. and J.S. Brownstein, *Putting the Public Back in Public Health - Surveying Symptoms of Covid-19*. N Engl J Med, 2020. **383**(7): p. e45.
22. Lapointe-Shaw, L., et al., *Web and phone-based COVID-19 syndromic surveillance in Canada: A cross-sectional study*. PLoS One, 2020. **15**(10): p. e0239886.
23. Chunara, R., Aman, S., Smolinski, M., & Brownstein, J. S., *Flu Near You: An Online Self-reported Influenza Surveillance System in the USA*. Online Journal of Public Health Informatics, 2013. **5**(1), e133.
24. Yousaf, A.R., et al., *A prospective cohort study in non-hospitalized household contacts with SARS-CoV-2 infection: symptom profiles and symptom change over time*. Clin Infect Dis, 2020.
25. CDC. *Disease Burden of Influenza*. 2020; Available from: <https://www.cdc.gov/flu/about/burden/index.html>.
26. 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, 2020. **69**(15): p. 458-464.
27. CDC. *COVID-NET Weekly Summary of U.S. COVID-19 Hospitalization Data*. Available from: https://gis.cdc.gov/grasp/COVIDNet/COVID19_3.html.
28. Kim, L., et al., *Risk Factors for Intensive Care Unit Admission and In-hospital Mortality among Hospitalized Adults Identified through the U.S. Coronavirus Disease 2019 (COVID-19)-Associated Hospitalization Surveillance Network (COVID-NET)*. Clin Infect Dis, 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, 2020. **69**(32): p. 1081-1088.
30. DHHS. *COVID-19 Estimated Patient Impact and Hospital Capacity by State*. Available from: <https://healthdata.gov/dataset/covid-19-estimated-patient-impact-and-hospital-capacity-state>.

31. CDC. *Nationwide Commercial Laboratory Seroprevalence Survey*. Available from: <https://covid.cdc.gov/covid-data-tracker/#national-lab>.
32. Reed, C., et al., *Estimating influenza disease burden from population-based surveillance data in the United States*. PLoS One, 2015. **10**(3): p. e0118369.
33. Rolfes, M.A., et al., *Annual estimates of the burden of seasonal influenza in the United States: A tool for strengthening influenza surveillance and preparedness*. Influenza Other Respir Viruses, 2018. **12**(1): p. 132-137.
34. CDC. *Estimated Influenza Illnesses, Medical visits, Hospitalizations, and Deaths in the United States — 2017–2018 influenza season*. 2019; Available from: <https://www.cdc.gov/flu/about/burden/2017-2018.htm>.
35. Madley-Dowd, P., et al., *The proportion of missing data should not be used to guide decisions on multiple imputation*. J Clin Epidemiol, 2019. **110**: p. 63-73.
36. Byambasuren, O., et al., *Estimating the extent of asymptomatic COVID-19 and its potential for community transmission: systematic review and meta-analysis*. medRxiv, 2020.

Accepted Manuscript

Figure Legends

Figure 1. Flow diagram of the methods to estimate total numbers of hospitalized, symptomatic illnesses, and infections from SARS-CoV-2 in the United States through adjustment of nationally reported case counts.

Accepted Manuscript

Table 1. Reported laboratory-confirmed coronavirus disease 2019 (COVID-19) cases and hospitalization status, by age and region, United States, February–September 2020.

	Reported cases ¹ N=	Rate of reported cases, per 100,000 population ^{1,2}	Percentage of reported cases hospitalized, median ^{1,3,4}	Rate of reported hospitalization, median, per 100,000 population ^{1,2,3}
Total	6,891,764	2,106	14%	296
Age group (years)				
0–4	109,317	552	6%	31
5–17	458,552	856	3%	26
18–49	3,974,817	2,877	7%	195
50–64	1,377,416	2,181	19%	418
≥ 65	971,662	1,853	43%	789
Health and Human Services (HHS) Region				
1 (CT, MA, ME, NH, RI, VT)	208,808	1,406	19%	271
2 (NJ, NY, PR, VI) ⁴	679,581	2,389	33%	786
3 (DE, DC, MD, PA, VA, WV)	468,065	1,518	14%	208
4 (AL, FL, GA, KY, MS, NC, SC, TN)	1,769,052	2,664	9%	241

5 (IL, IN, MI, MN, OH, WI)	882,355	1,679	15%	258
6 (AR, LA, NM, OK, TX)	1,092,515	2,576	12%	312
7 (IA, KS, MO, NE)	308,636	2,185	8%	174
8 (CO, MT, ND, SD, UT, WY)	200,390	1,651	7%	123
9 (AZ, CA, HI, NV, AS, MP, FSM, GU, RMI, PW)	1,117,180	2,183	14%	310
10 (AK, ID, OR, WA)	165,182	1,162	8%	92

¹Patient age imputed if missing (17% of cases).

² Population estimated using CDC Wonder Bridged-Race estimates[16].

³ Patient hospitalization status imputed if missing (86% of cases).

⁴ For hospitalization imputation, the regional proportion of cases reported as hospitalized in region 2 was estimated excluding NYC due to large discrepancy between national and jurisdiction reports.

Table 2. Sources of under-detection included in model-based estimates of the incidence of COVID-19, United States, February–September 2020.

	Parameter	Data source [reference]	Observed value	Statistical distribution included in model
Hospitalized				
A	SARS-CoV-2 assay sensitivity	Systematic Review[17]	2%–21% false negative rate across included studies	Uniform (0.79, 0.98) (Same values as non-hospitalized)
B	SARS-CoV-2 test ordered and completed	IBM Watson and COVID Near You	Median (range): 0–17 yrs 33% (15%–55%) 18–49 yrs 50% (21%–96%) 50–64 yrs 51% (18%–97%) ≥65 yrs 54% (6%–98%)	Beta PERT, varies by age and date of case report (see Supplemental Table 1) (Values specific to hospitalized settings)
Non-hospitalized				
A	SARS-CoV-2 assay sensitivity	Systematic Review [17]	2%–21% false negative rate across included studies	Uniform (0.79, 0.98) (Same values as hospitalized)
B	SARS-CoV-2 test ordered and completed	IBM Watson and COVID Near You	Median (range): 0–17 yrs: 43% (1%–71%) 18–49 yrs: 53% (6%–99%)	Beta PERT, varies by age and date of case report (see Supplemental Table 1) (Values specific to outpatient)

			50–64 yrs: 58% (6%–98%)	settings)
			≥65 yrs: 54% (6%–99%)	
C	Symptomatic patient seeks care	COVID Near You and Flu Near You	Median (range): 0–17 yrs: 26% (13%–49%) 18–49 yrs: 34% (15%–65%) 50–64 yrs: 35% (13%–55%) ≥65 yrs: 40% (11%–60%)	Beta PERT, varies by age and HHS region (see Supplemental Table 2) (Values specific to outpatient settings)
Total infections				
D	Infected person is asymptomatic	[36], [8]	0–64 yrs: 5%–24% ≥65 yrs: 5%–32%	Uniform, varies by age

Table 3. Estimates of hospitalized persons with COVID–19 and rates per 100,000 population, United States, February–September, 2020.

	Estimated Hospitalizations	95% UI ¹	Rate, per 100,000 ²	95% UI ¹
Overall³	2,397,777	2,053,156–2,855,843	733	628–873
Age group (years)				
0–4	20,719	16,595–26,069	105	84–132

5–17	44,321	33,300–58,552	83	62–109
18–49	652,741	530,955–823,453	472	384–596
50–64	642,358	538,092–778,266	1,017	852–1,232
≥ 65	1,022,295	826,438–1,361,730	1,950	1,576–2,597

HHS Region

1 (CT, MA, ME, NH, RI, VT)	103,347	86,983–125,978	696	586–848
2 (NJ, NY, PR, VI) ⁴	694,079	580,828–878,399	2,440	2,042–3,087
3 (DE, DC, MD, PA, VA, WV)	152,597	129,196–181,597	495	419–589
4 (AL, FL, GA, KY, MS, NC, SC, TN)	349,780	289,336–423,126	527	436–637
5 (IL, IN, MI, MN, OH, WI)	330,948	278,985–395,956	630	531–754
6 (AR, LA, NM, OK, TX)	288,441	231,390–357,417	680	546–843
7 (IA, KS, MO, NE)	55,692	45,600–68,225	394	323–483
8 (CO, MT, ND, SD, UT, WY)	39,413	33,449–47,559	325	276–392
9 (AZ, CA, HI, NV, AS, MP, FSM, GU, RMI, PW)	347,069	291,807–414,925	678	570–811
10 (AK, ID, OR, WA)	33,552	28,430–41,594	236	200–293

¹Uncertainty interval (UI)

²Population estimated using CDC Wonder Bridged-Race estimates[16]

³ Due to rounding, age group and region estimates may not sum to overall estimates

⁴ For hospitalization imputation, the regional proportion of cases reported as hospitalized in region 2 was estimated excluding NYC due to large discrepancy between national and jurisdiction reports.

Accepted Manuscript

Table 4. Estimates of symptomatic illnesses from SARS-CoV-2 infection and rates per 100,000 population, United States, February–September, 2020.

	Estimated symptomatic illnesses	95% UI ¹	Rate, per 100,000 ²	95% UI
Overall³	44,769,417	36,920,353–55,535,659	13,684	11,285–16,975
Age group (years)				
0–4	1,144,532	903,194–1,500,315	5,777	4,559–7,573
5–17	4,719,785	3,722,649–6,177,733	8,807	6,947–11,528
18–49	25,096,725	19,137,381–34,524,124	18,162	13,850–24,985
50–64	8,926,318	6,873,250–11,928,318	14,133	10,883–18,887
≥ 65	4,556,384	3,569,223–6,172,269	8,690	6,807–11,772
HHS Region				
1 (CT, MA, ME, NH, RI, VT)	1,613,724	1,321,058–2,018,578	10,864	8,894–13,590
2 (NJ, NY, PR, VI)	7,597,800	5,828,412–10,470,229	26,705	20,486–36,801
3 (DE, DC, MD, PA, VA, WV)	3,179,080	2,627,466–3,915,693	10,307	8,519–12,696
4 (AL, FL, GA, KY, MS, NC, SC, TN)	10,263,209	8,428,418–12,805,689	15,457	12,694–19,286

5 (IL, IN, MI, MN, OH, WI)	5,351,673	4,327,360–6,904,494	10,185	8,236–13,141
6 (AR, LA, NM, OK, TX)	6,200,307	5,074,596–7,815,354	14,618	11,964–18,426
7 (IA, KS, MO, NE)	1,796,811	1,456,136–2,307,574	12,722	10,310–16,339
8 (CO, MT, ND, SD, UT, WY)	1,329,434	1,087,300–1,676,711	10,952	8,957–13,813
9 (AZ, CA, HI, NV, AS, MP, FSM, GU, RMI, PW)	6,155,322	5,093,684–7,539,673	12,026	9,952–14,731
10 (AK, ID, OR, WA)	1,177,495	956,766–1,476,146	8,282	6,729–10,382

¹Uncertainty interval (UI)

²Population estimated using CDC Wonder Bridged-Race estimates[16]

³Due to rounding, age group and region estimates may not sum to overall estimates

Table 5. Estimates of total infections and rates per 100,000 population, United States, February–September, 2020.

	Estimated total infections	95% UI ¹	Rate, per 100,000 ²	95% UI
Overall³	52,885,526	42,527,569–66,810,205	16,165	12,999–20,421
Age group (years)				
0–4	1,342,212	1,022,465–1,811,583	6,775	5,161–9,145
5–17	5,538,766	4,222,053–7,451,900	10,336	7,879–13,906
18–49	29,421,481	21,798,393–41,330,693	21,292	15,775–29,911
50–64	10,484,802	7,860,849–14,346,364	16,601	12,446–22,715
≥ 65	5,636,607	4,139,528–8,024,420	10,750	7,895–15,305
HHS Region				
1 (CT, MA, ME, NH, RI, VT)	1,910,156	1,530,556–2,427,485	12,860	10,304–16,343
2 (NJ, NY, PR, VI)	8,977,706	6,780,805–12,534,610	31,555	23,834–44,057
3 (DE, DC, MD, PA, VA, WV)	3,759,656	3,025,633–4,721,730	12,190	9,810–15,309
4 (AL, FL, GA, KY, MS, NC, SC, TN)	12,107,021	9,741,202–15,448,374	18,234	14,671–23,266
5 (IL, IN, MI, MN, OH, WI)	6,324,790	5,002,112–8,272,471	12,037	9,520–15,744

6 (AR, LA, NM, OK, TX)	7,315,403	5,856,636–9,423,602	17,248	13,808–22,218
7 (IA, KS, MO, NE)	2,122,340	1,684,158–2,788,524	15,027	11,925–19,744
8 (CO, MT, ND, SD, UT, WY)	1,569,175	1,250,443–2,013,970	12,927	10,301–16,591
9 (AZ, CA, HI, NV, AS, MP, FSM, GU, RMI, PW)	7,243,925	5,876,211–9,098,261	14,153	11,481–17,776
10 (AK, ID, OR, WA)	1,391,488	1,106,862–1,775,644	9,787	7,785–12,489

¹Uncertainty interval (UI)

²Population estimated using CDC Wonder Bridged-Race estimates [16]

³Due to rounding, age group and region estimates may not sum to overall estimates

Figure 1

