

How many SARS-CoV-2 infected people require hospitalization? Using random sample testing to better inform preparedness efforts

Nir Menachemi, PhD, MPH

Indiana University Richard M. Fairbanks School of Public Health and Regenstrief Institute, Inc.
nirmen@iu.edu

Brian E. Dixon, PhD

Indiana University Richard M. Fairbanks School of Public Health and Regenstrief Institute, Inc.
bedixon@regenstrief.org

Kara K. Wools-Kaloustian, MD

Indiana University School of Medicine
kwools@iu.edu

Constantin Yiannoutsos, PhD

Indiana University Richard M. Fairbanks School of Public Health
cyiannou@iu.edu

Paul K. Halverson, DrPH

Indiana University Richard M. Fairbanks School of Public Health
pkhalver@iu.edu

Word count: 1,995 (not including abstract or references)

Funding: This study was funded by the Indiana State Department of Health

Corresponding Author: Nir Menachemi, PhD, MPH Fairbanks Endowed Chair, Professor, and Department Chair Health Policy and Management, nirmena@iu.edu, 1050 Wishard Blvd. Suite 5130 Indianapolis IN, 46202. (317) 274-0388

Keywords: hospitalization rate, COVID-19, case-hospitalization ratio, infection-hospitalization ratio

This is the author's manuscript of the article published in final edited form as:

Menachemi, N., Dixon, B. E., Wools-Kaloustian, K. K., Yiannoutsos, C. T., & Halverson, P. K. (2021). How Many SARS-CoV-2–Infected People Require Hospitalization? Using Random Sample Testing to Better Inform Preparedness Efforts. *Journal of Public Health Management and Practice*, 27(3), 246–250. <https://doi.org/10.1097/PHH.0000000000001331>

Abstract

Context: Existing hospitalization ratios for COVID-19 typically use case counts in the denominator which problematically underestimates total infections because asymptomatic and mildly infected persons rarely get tested. As a result, surge models that rely on case counts to forecast hospital demand may be inaccurately influencing policy and decision-maker action.

Objective: Based on SARS-CoV-2 prevalence data derived from a statewide random sample (as opposed to relying on reported case counts), we determine the infection-hospitalization ratio (IHR), defined as the percentage of infected individuals who are hospitalized, for various demographic groups in Indiana. Further, for comparison, we show the extent to which case-based hospitalization ratios, compared to the IHR, overestimate the probability of hospitalization by demographic group.

Design: Secondary analysis of statewide prevalence data from Indiana, COVID-19 hospitalization data extracted from a statewide health information exchange, and all reported COVID-19 cases to the state health department.

Setting: State of Indiana as of April 30, 2020.

Main Outcome Measure(s): demographic-stratified Infection-hospitalization ratios and case-hospitalization ratios.

Results: The overall IHR was 2.1% and varied more by age than by race or sex. IHR estimates ranged from 0.4% for those under 40 years to 9.2% for those over 60 years. Hospitalization rates based on case counts overestimated the IHR by a factor of ten, but this overestimation differed by demographic groups, especially age.

Conclusions: In this first study of the IHR based on population prevalence, our results can improve forecasting models of hospital demand—especially in preparation for the upcoming winter period when an increase in SARS CoV-2 infections is expected.

Introduction

U.S. hospitalization ratios for SARS-CoV-2, are typically calculated as the proportion of known cases, based on confirmatory laboratory testing, that are admitted to a given health system. These ratios range from 29% to 40%, depending on the characteristics of the population studied (1, 2). Among those hospitalized, 14%-30% required intensive care (1, 3) of whom a high percentage (86%) required mechanical ventilation (3). Risk factors for COVID-19 hospitalization include older age (3-5), having underlying chronic conditions (e.g. hypertension, obesity, chronic lung disease) (4, 5), and positive smoking status (5).

Calculating hospitalization ratios based on diagnosed cases is problematic, because asymptomatic and mildly symptomatic infected individuals rarely get tested and frequently go undiagnosed. Moreover, the number of cases reported is biased by the availability of tests and the motivation for testing both of which can fluctuate over time. As such, cases underestimate the number of infections by a factor of six to 24 times (6). Such inaccuracies could adversely affect prediction models of hospital demand, which could lead to either over or under-preparedness of staff, supplies, or equipment— all of which is in short supply during a pandemic. In addition, without accurate data on hospitalization ratios by age, race, or sex, decision-makers are unable to tailor decisions to the unique risks of specific, and often vulnerable, populations.

Unlike the case-hospitalization ratio, which is the percent of diagnosed cases that are hospitalized, the infection-hospitalization ratio (IHR), defined as the percentage of infected persons who are hospitalized, uses population prevalence information in the denominator.

Calculating an IHR has been previously challenging because of the lack of reliable information about the population prevalence of infections in a given a state. Using SARS-CoV-2 prevalence data derived from the first statewide random sample study in the US (7), we present demographically stratified IHR estimates from Indiana. Because random population testing also identifies asymptomatic and mildly infected persons, the denominator used in the current study reduces the bias inherent in calculations based on diagnosed cases only. In the current study, we calculate both the IHR and the case-hospitalization ratio for each demographic group and show how the discrepancy between the two calculations varies by age, race, and sex.

Methods

We analyze statewide data from Indiana including case counts, hospitalizations, and estimated infections as of April 30, 2020. Data on confirmed SARS-CoV-2 cases were obtained from the state health department's official public COVID-19 dashboard (8). Data on hospitalizations were obtained from the Indiana Network for Patient Care, the nation's oldest and most comprehensive health information exchange (9). These data represent the best information on the unique number of patients hospitalized in the state for COVID-19 and differ from the daily aggregate hospitalized case counts available on the state's dashboard. Data on the estimated number of infections were obtained from the aforementioned prevalence study that assessed both active infections and seroprevalence among a randomly selected sample of state residents (7). Briefly, the prevalence study utilized a list of state residents derived from tax records to randomly select persons aged ≥ 12 years from within Indiana's 10 public health preparedness districts. Institutionalized individuals were excluded and prevalence estimates

were adjusted for nonresponse based on age, race, and ethnicity. Based on hospital data availability, this study focuses on adults aged 18 and older.

We calculate the IHR for all hospitalizations, and for ICU admissions (heretofore ICU-IHR), by dividing the number of hospitalizations, or ICU stays, by the estimated number of state residents that were cumulatively infected with SARS-CoV-2, using data from the prevalence study, as of the end of April 2020. Cumulative infections were determined by multiplying the proportion of randomly selected individuals that either showed evidence of current viral infection or the presence of antibodies for SARS-CoV-2 by state population estimates obtained from the US Census Bureau. Importantly, because the population prevalence, used as the denominator in the IHR, was determined based on a random sample (and not merely based on reported cases), the infections we identify represent all people who are infected (including asymptomatic and mildly symptomatic individuals). The vast majority of these people are not expected to have severe outcomes. Thus, this group represents the ideal denominator needed to calculate the IHR.

Demographic groups included three categories of age (18– 39 years, 40-59 years, and 60 years or greater), race measured dichotomously (white or non-white), and sex (male or female). IHRs were calculated as the number of hospitalizations divided by the estimated number of infections plus the number of previous COVID-19 deaths as of April 30, 2020. The addition of COVID-19 deaths in the denominator ensures that all previous infections in the state are captured in the IHR estimate. Case-hospitalizations ratios were calculated as the number of hospitalizations divided by the number of reported cases. Both the ICU-IHR and the case-ICU admission ratios were also calculated using statewide ICU admission data. Lastly, to determine

the magnitude that cases, relative to infections, overestimate the hospitalization rate, we divide the case-hospitalization ratio by the IHR for each demographic group.

Results

As of April 30, 2020, there were 4,006 adult hospitalizations for COVID-19 in Indiana including 953 ICU admissions (see Table 1). As of this date, 18,422 confirmed cases had been reported to the state health department, one-tenth of the estimated 187,829 infections (95% CI: 134,644–249,092) in the state, based on the random sample testing study. Based on the number of infections, the statewide population prevalence was 2.8%. Hospitalizations and ICU admissions, but not cases or infections, were concentrated among those 60 years and older.

In Table 2, we present the IHR and ICU-IHR for the overall population and by demographic groups. For comparison, we also present the case-hospitalization rates for overall hospitalizations and ICU admissions. The overall IHR was 2.1% and increased with age from 0.4% for the youngest age group (18-39 year-olds) to 9.2% for the oldest age group (60 years and older). The overall case-hospitalization rate was 21.8% (10.4 times greater than the comparable IHR) and also increased with age from 8.4% for the youngest group to 35.0% among those in the oldest group.

The ICU-IHR was 0.5% and ranged from 0.05% in the youngest age group to 2.6% in the oldest age group. The ICU case-hospitalization rate was 10.2 times greater than the comparable IHR and ranged from 1.2% in the youngest group to 9.8% for the oldest group. Unlike age, much smaller and nonsignificant variations in ratios were observed for race and sex groups.

Discussion

In this first study to calculate the IHR based on population prevalence estimates, we found that the probability of hospitalization given infection was 2.1%. This suggests that one in approximately 48 SARS-CoV-2 infections led to an inpatient stay. More importantly, the IHR was highest among those over 60 years where an estimated 1 in 11 infections resulted in a hospitalization. Approximately 1 in 200 overall infections resulted in an ICU admission with a significantly higher rate (1 in 38) among those over 60 years of age.

To date, COVID-19 surge planning has been done without enough specific information about the variability in the IHR by demographic characteristics. Early forecasters trying to understand the impact of the novel coronavirus were forced to make many assumptions (11, 12) that were subsequently refuted by accumulating data (13). Many forecasting models assumed a previously defined fraction of the infected population would require hospitalization without being able to consider the specific differences in the IHR by age, sex, and race (14). Using static estimates of projected hospitalizations can distort the true impact of COVID-19 when infection rates differ by demographic characteristics. This can create a false sense of security among public health officials and hospital administrators (e.g., when infection rates are truly higher for elderly populations but case counts do not reflect this sufficiently) or result in over-preparation of health care resources (e.g., when infection rates are higher in younger groups), causing the unnecessary scale back of operations in preparation for crowds who never materialize for COVID-19 treatment. We believe the relatively low US hospitalization utilization for COVID-19 in some summer months of 2020 was due to the latter situation. By November 2020, when hospitalization rates increased, it was likely due to increases in infections among

the highest-risk groups (consistent with the former example). The current study provides generalizable population estimates that can improve forecasting of hospital demand—especially in preparation for the upcoming periods when an increase in SARS CoV-2 infections is expected. Knowing, for example, that individuals over 60, compared to those under 40, are 23 times more likely to need hospital care if infected with SARS-CoV-2 is important for policy and planning considerations.

We presented hospitalization rates based on case counts (which are typically limited to diagnosed cases) to show how they overestimate the IHR. Overall, in Indiana, hospitalization rates based on cases overestimated the infection-hospitalization ratio by 10-fold with magnitudes that differed most by age. In the oldest age group, case counts were most similar to the IHR but still overestimated the hospitalization rate by more than three-fold. This smaller overestimation is likely due to higher clinical testing rates in this age group, which is known to be at the highest risk of symptomatic disease, hospitalizations, and poor outcomes (e.g., ICU admission, death). Moreover, nursing homes residents, who are concentrated in the oldest age group, are known to be at higher risk and may be tested more often (10). Higher testing rates help case counts more closely resemble infections.

Public health officials as well as hospital administrators can apply the knowledge gained from the current study to their current COVID-19 response efforts. For example, several states followed Indiana and conducted regional or statewide prevalence studies since April 2020. Using our methods, these states could calculate IHR rates for their populations using hospitalizations reported to state health agencies. All states, including those that do not conduct prevalence studies can use the IHR values presented herein to better understand the

relationship between infections and hospital resource demands. For example, 100 cases among individuals 60+ years would likely result in at least 11 hospitalizations with the upper bound estimate equaling the extent to which cases underestimate infections (which in Indiana was 3.8 times in this age group). Despite being the first study to estimate the IHR based on a statewide random sample, our study has several limitations worth mentioning. First, although nonresponse was accounted for in the original random sample using standard statistical approaches, response bias cannot be ruled out from having affected the population prevalence estimates. Moreover, data availability did not allow us to calculate the IHR in persons under 18 so generalizability to younger age groups is not possible based upon the current results. Third, by design, the IHR is a population-based metric that is not intended to describe individual risks but instead elucidate the probability of needing hospital services given infection. The IHR is more suited for policy and planning purposes than clinical decision making. Fourth, we recognize that race data in particular may not be accurately captured in hospitalization records. The degree to which this affects the results we present is unknown. Moreover, due to statistical power and sampling issues, race was dichotomized in the Menachemi et al (2020) prevalence study resulting in the binary measure of race used in the current study. We recognize that this is a limitation because the experiences of minorities, especially with SARS-CoV-2, may not be uniform. Lastly, our study was based in Indiana at a given point in time which may affect the generalizability of our findings. Nevertheless, in the absence of COVID-19 outpatient therapeutics, the risk of hospitalization given infection should be similar in other similar populations.

In conclusion, our study provides scientific precision regarding the proportion of SARS-CoV-2 infected individuals that require hospital services. With these data, surge planning efforts and COVID-19 models can be used to more reliably prepare hospitals for the upcoming fall and winter seasons. Lastly, our data show how case-counts lead to overestimated hospitalization rates, especially in groups less likely to be tested.

Implications for Policy & Practice

- In order to forecast demand for hospital services during the pandemic, public health researchers have used hospitalization ratios based on case-counts. Case counts are problematic because they are known to underestimate infections due to excluding most asymptomatic and mildly infected individuals.
- Based on the first statewide random sample study of SARS-CoV-2 prevalence, we were able to calculate the infection-hospitalization ratio (IHR), both overall and by demographic group, to determine the probability of hospitalization given infection. IHRs use all infections in the denominator to calculate the hospitalization ratio which provides more reliable information than relying on cases counts.
- We show how hospitalization ratios based on case-counts differentially overestimate the IHR by demographic groups especially age.
- Our results can improve forecasting models designed to predict COVID-19 hospital resource needs (e.g., staff, equipment, supplies) especially as infections increase.

Table 1: Statewide counts of COVID-19 cases, hospitalizations, and estimated infections in Indiana as of April 30, 2020.

	No. of estimated infections ¹ (95% CI)	No. of Reported Cases	No. of COVID-19 Hospitalizations (non-ICU and ICU)	No. of COVID-19 ICU Admissions	No. of Deaths
Age Category					
18 – 39 years	107,588 (67,022 to 151,681)	5,080	424	59	14
40 – 59 years	52,973 (32,054 to 84,352)	6,644	1,235	239	81
60+ years	24,377 (14,774 to 35,457)	6,698	2,347	655	1,004
Race					
White	154,141 (97,052 to 188,394)	11,392	2,444	671	715
Nonwhite	49,425 (17,396 to 97,213)	4,989	1,328	248	384
Sex					
Male	108,199 (63,061 to 165,949)	8,341	2,071	527	493
Female	82,259 (54,612 to 112,637)	10,004	1,935	426	580
Total	187,829 (134,644 to 249,092)	18,422	4,006	953	1,099

Note: ICU is intensive care unit

¹Based on a random sample testing study, see Menachemi et al. (2020)

Table 2: Case-Hospitalization and Infection-Hospitalization Ratios in Indiana as of April 30, 2020

	Case-Hospitalization Ratio¹	Case-ICU Ratio²	Infection-Hospitalization Ratio³ (95% CI)	Infection-ICU Ratio⁴ (95% CI)	Magnitude that Cases, Relative to Infections, Overestimate the Hospitalization Rate⁵
Age Category					
18 – 39 years	8.4%	1.2%	0.4% (0.3 – 0.6)	0.05% (.04 – .09)	21.0x
40 – 59 years	18.6%	3.6%	2.3% (1.5 – 3.8)	0.45% (.28 – .74)	8.1x
60+ years	35.0%	9.8%	9.2% (6.4 – 14.9)	2.6% (1.8 – 4.2)	3.8x
Race					
White	21.5%	5.9%	1.7% (1.3 – 2.5)	0.47% (.35 – .69)	12.6x
Nonwhite	26.6%	5.0%	3.2% (1.4 – 7.5)	0.59% (.25 – 1.4)	8.3x
Sex					
Male	24.8%	6.3%	1.9% (1.2 – 3.3)	0.49% (.32 – .83)	13.1x
Female	19.3%	4.3%	2.3% (1.7 – 3.5)	0.52% (.38 – .77)	8.4x
Total	21.8%	5.2%	2.1% (1.6 – 3.0)	0.51% (.38 – .70)	10.4x

Note: ICU is intensive care unit

¹Calculated as number of hospitalizations divided by the number of reported cases

²Calculated as number of ICU stays divided by the number of reported cases

³Calculated as number of hospitalizations divided by the estimated number of infections plus the number of COVID-19 deaths

⁴Calculated as number of ICU stays divided by the estimated number of infections plus the number of COVID-19 deaths

⁵Case-Hospitalization Ratio divided by Infection-Hospitalization Ratio

Literature Cited

1. Myers LC, Parodi SM, Escobar GJ, Liu VX. Characteristics of Hospitalized Adults With COVID-19 in an Integrated Health Care System in California. *JAMA*. 2020;323(21):2195-8.
2. Price-Haywood EG, Burton J, Fort D, Seoane L. Hospitalization and Mortality among Black Patients and White Patients with Covid-19. *New England Journal of Medicine*. 2020;382(26):2534-43.
3. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting Characteristics, Comorbidities, and Outcomes Among 5700 Patients Hospitalized With COVID-19 in the New York City Area. *JAMA*. 2020;323(20):2052-9.
4. Garg S, Kim L, Whitaker M, O'Halloran A, Cummings C, Holstein R, 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):458-64.
5. Killerby ME, Link-Gelles R, Haight SC, Schrodt CA, England L, Gomes DJ, et al. Characteristics Associated with Hospitalization Among Patients with COVID-19 - Metropolitan Atlanta, Georgia, March-April 2020. *MMWR Morbidity and mortality weekly report*. 2020;69(25):790-4.
6. Havers FP, Reed C, Lim T, Montgomery JM, Klena JD, Hall AJ, et al. Seroprevalence of Antibodies to SARS-CoV-2 in 10 Sites in the United States, March 23-May 12, 2020. *JAMA Internal Medicine*. 2020.
7. Menachemi N, Yiannoutsos CT, Dixon BE, Duszynski TJ, Fadel WF, Wools-Kaloustian KK, et al. Population Point Prevalence of SARS-CoV-2 Infection Based on a Statewide Random Sample - Indiana, April 25-29, 2020. *MMWR Morb Mortal Wkly Rep*. 2020;69(29):960-4.
8. Indiana COVID-19 dashboard Indianapolis, IN: Indiana State Department of Health; 2020 [Available from: <https://www.coronavirus.in.gov/2393.htm>].
9. McDonald CJ, Overhage JM, Barnes M, Schadow G, Blevins L, Dexter PR, et al. The Indiana network for patient care: a working local health information infrastructure. An example of a working infrastructure collaboration that links data from five health systems and hundreds of millions of entries. *Health Aff (Millwood)*. 2005;24(5):1214-20.
10. D'Adamo H, Yoshikawa T, Ouslander JG. Coronavirus Disease 2019 in Geriatrics and Long-Term Care: The ABCDs of COVID-19. *J Am Geriatr Soc*. 2020;68(5):912-7.
11. Adelman D. Thousands Of Lives Could Be Saved In The US During The COVID-19 Pandemic If States Exchanged Ventilators. *Health Affairs*. 2020;39(7):1247-52.

12. Bartsch SM, Ferguson MC, McKinnell JA, O'Shea KJ, Wedlock PT, Siegmund SS, et al. The Potential Health Care Costs And Resource Use Associated With COVID-19 In The United States. *Health Affairs*. 2020;39(6):927-35.
13. Jewell NP, Lewnard JA, Jewell BL. Caution Warranted: Using the Institute for Health Metrics and Evaluation Model for Predicting the Course of the COVID-19 Pandemic. *Annals of Internal Medicine*. 2020;173(3):226-7.
14. CDC. COVID-19 Forecasts: Hospitalizations 2020 [Available from: <https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/hospitalizations-forecasts.html>].