



Staffing Levels and COVID-19 Cases and Outbreaks in U.S. Nursing Homes

Rebecca J. Gorges, PhD*   and R. Tamara Konetzka, PhD*[†]

BACKGROUND/OBJECTIVES: Nursing homes have experienced a disproportionate share of COVID-19 cases and deaths. Early analyses indicated that baseline quality was not predictive of nursing home cases, but a more nuanced study of the role of nurse staffing is needed to target resources and better respond to future outbreaks. We sought to understand whether baseline nurse staffing is associated with the presence of COVID-19 in nursing homes and whether staffing impacts outbreak severity.

DESIGN: We analyzed Centers for Medicare & Medicaid Services (CMS) facility-level data on COVID-19 cases and deaths merged with nursing home and county characteristics. We used logistic regressions to examine the associations of staffing levels from Nursing Home Compare with the outcomes of any COVID-19 cases and, conditional on at least one case, an outbreak. Among facilities with at least one case, we modeled count of deaths using hurdle negative binomial-2 regressions.

SETTING: All nursing homes in the CMS COVID-19 Nursing Home Dataset with reports that passed the CMS Quality Assurance Check as of June 25, 2020.

PARTICIPANTS: Residents of nursing homes that met COVID-19 reporting requirements.

MEASUREMENTS: A nursing home is defined as having at least one case if one or more confirmed or suspected COVID-19 case among residents or staff is reported. Conditional on at least one case, we examine two outcomes: an outbreak, defined as confirmed cases/certified beds >10% or total confirmed and suspected cases/beds >20% or >10

deaths, and the total number of deaths attributed to COVID-19 among residents and staff.

RESULTS: A total of 71% of the 13,167 nursing homes that reported COVID-19 data as of June 14 had at least one case among residents and/or staff. Of those, 27% experienced an outbreak. Higher registered nurse-hours are associated with a higher probability of experiencing any cases. However, among facilities with at least one case, higher nurse aide (NA) hours and total nursing hours are associated with a lower probability of experiencing an outbreak and with fewer deaths. The strongest predictor of cases and outbreaks in nursing homes is per capita cases in the county.

CONCLUSION: The prevalence of COVID-19 in the community remains the strongest predictor of COVID-19 cases and deaths in nursing homes, but higher NA hours and total nursing hours may help contain the number of cases and deaths. *J Am Geriatr Soc* 00:1-5, 2020.

Keywords: COVID-19; nursing homes; long-term care

From the *Department of Public Health Sciences, The University of Chicago, Chicago, Illinois; and the [†]Department of Medicine, The University of Chicago, Chicago, Illinois.

Address correspondence to Rebecca Gorges, PhD, Department of Public Health Sciences, The University of Chicago, 5841 S. Maryland Avenue, MC2000, Chicago, IL 60637. E-mail: rjgorges@uchicago.edu; Twitter: [@gorges_rebecca](https://twitter.com/gorges_rebecca)

This article was published online on 28 August 2020. An error was subsequently identified in the author name. This notice is included in both the online and print versions to indicate that they have been corrected on 3 September 2020.

DOI: 10.1111/jgs.16787

Residents and staff at nursing homes to date account for more than 282,000 cases and 54,000 deaths, or 11% of cases and 43% of deaths from COVID-19 in the United States.¹ In some ways, the high rates are not surprising. Nursing homes house, in close quarters, large numbers of individuals with multiple comorbidities who need hours of hands-on care every day. These facilities are often understaffed and depend on Medicaid reimbursement for most of their residents. During the pandemic, they have had to compete with hospitals for personal protective equipment (PPE), still in short supply in many areas, putting both staff and residents at risk. These circumstances lend an aura of inevitability to cases of COVID-19 appearing in nursing homes.

A natural question is whether the risk can be identified and mitigated. Recent examinations of selected states found that the percentage of non-White residents was a strong predictor of cases.² In addition to race, larger facility size,

urban location, non-chain status, state,³ and county-level virus prevalence⁴ were associated with more cases. None of these studies found an association between cases and Nursing Home Compare (NHC) star ratings, implying that location of the nursing home, rather than quality level, determined vulnerability to COVID-19. However, it seems counterintuitive that nursing home staffing, a primary driver of quality, plays no role. Single-state studies of nursing homes in California⁵ and Connecticut⁶ found that facilities with lower registered nurse (RN) staffing have more cases. Our objective is to take a nuanced look at the role of nursing home staffing in COVID-19 cases and deaths using national data. First, are staffing levels associated with the probability of any COVID-19 cases in a facility? Second, among facilities experiencing at least one case, is staffing associated with severity of the outbreak that develops?

METHODS

Study Sample and Data Sources

We merged nursing home COVID-19 case and death counts with facility and county characteristics. COVID-19 cases and deaths are from the COVID-19 Nursing Home Dataset⁷ released by the Centers for Medicare & Medicaid Services (CMS) on June 25, 2020. This file contains weekly

case and death counts among residents and staff reported by facilities to CMS. We then linked nursing homes by provider ID to two sources of facility-level data: NHC archives and LTCFocus.⁸ NHC provides the nursing home location, ownership, number of beds, NHC star ratings, and staffing levels as of the first quarter 2020. Using staffing levels from early 2020 reflects the extent to which nursing homes were sufficiently staffed directly before the pandemic that could enable better infection control once the pandemic started. LTCFocus contains information on resident characteristics (percentage Medicaid, percentage White) as of 2017. Finally, we include two county-level characteristics: COVID-19 cases per 1,000 residents as of June 14, 2020,⁹ and metropolitan status (U.S. Department of Agriculture Rural Urban Continuum Codes).

The study sample consists of 13,167 nursing facilities, the 85% of facilities in NHC that reported COVID-19 data for at least one of the first 4 weeks of CMS reporting and have staffing hours in NHC.

Outcomes

We examined two main outcomes: (1) whether facilities had any COVID-19 cases, and (2) among facilities with at least one case, the size of the outbreak. We defined any cases as a binary outcome indicating whether the facility

Table 1. Sample Characteristics by Any Cases and Outbreak

	Any cases			Among at least one case, outbreak		
	Yes	No	P value	Yes	No	P value
N	9,352	3,815		2,569	6,783	
No. of certified beds, mean	116.0	88.5	<.001	133.2	109.5	<.001
For profit, %	69.6	70.9	<.001	75.3	67.5	<.001
Government, %	5.6	7.7		3.5	6.4	
Nonprofit, %	24.8	21.5		21.2	26.1	
Chain, %	59.7	59.6	.925	57.5	60.5	.011
Residents, Medicaid, %	59.5	60.0	.289	61.5	58.7	<.001
Residents, White race, %	79.0	84.3	<.001	74.3	80.7	<.001
County-level confirmed cases/1,000 residents, mean	6.0	2.9	<.001	8.8	4.9	<.001
Metropolitan county, %	77.0	60.3	<.001	88.5	72.6	<.001
Resident access to testing in facility, %	98.2	95.2	<.001	99.0	97.9	<.001
<1-wk supply of one or more PPE type, %	22.1	17.8	<.001	21.7	22.2	.591
Shortage of nursing staff, %	15.6	12.9	<.001	18.8	14.4	<.001
Shortage of clinical staff, %	3.0	2.1	.003	3.7	2.7	.013
Shortage of aides, %	17.4	15.9	.042	19.6	16.5	<.001
Shortage of other staff, %	9.7	7.4	<.001	12.2	8.8	<.001
Overall rating, mean	3.2	3.2	.699	3.1	3.2	<.001
Health inspection rating, mean	2.8	2.9	<.001	2.7	2.8	<.001
QM rating, mean	3.7	3.5	.000	3.8	3.7	.094
Staffing rating, mean	3.1	3.0	.093	2.9	3.1	.000
Adjusted nurse aide hours, mean	2.3	2.3	.921	2.2	2.4	.000
Adjusted LPN hours, mean	.9	.9	.082	.9	.9	.155
Adjusted RN hours, mean	.7	.7	.660	.7	.7	.000
Adjusted total nurse hours, mean	3.9	3.9	.691	3.8	3.9	.000
Adjusted RN/total nurse hours, mean	17.9	17.8	.433	17.6	18.0	.032

Note: Facility must have reported data and data meets Centers for Disease Control and Prevention (CDC) quality check criteria for at least 1 week.

COVID data source: CMS/CDC Nursing Facility data, June 25, 2020.

Definition of outbreak: >10% confirmed cases/beds or >20% confirmed plus suspected cases/beds or ≥10 deaths.

P value from chi-square test for binary/categorical measures and *t* test for continuous measures.

reported at least one confirmed or suspected COVID-19 case among residents or staff. We measured outbreak size in two ways. First, we determined whether the facility met the CMS threshold of an outbreak: greater than one confirmed case per 10 certified beds or greater than one total confirmed and suspected cases per five certified beds or greater than 10 deaths.¹⁰ Second, we examined the number of COVID-19 deaths as a count variable including suspected and confirmed deaths among residents and staff.

Staffing Measures

Staffing levels are from NHC and are derived from the Payroll-Based Journal. We used the case-mix adjusted hours per resident per day for nursing aides (NAs), licensed practical nurses (LPNs), and RNs. We also created a measure to characterize the intensity of RN staffing: share of total nursing hours provided by RNs.

Statistical Analysis

We began by examining unadjusted associations between facility characteristics and county characteristics with the two outcomes of any cases and, among facilities with at least one case, the presence of an outbreak. Next, we used multivariable regressions to determine the adjusted association between staffing levels and COVID-19 cases/outbreaks. We conducted two sets of regressions for each outcome, one examining the individual effects of NA, LPN, and RN

hours and one examining the effect of total nursing hours and RN hours/total nursing hours. Staffing measures were converted from continuous measures of hours to categorical variables by dividing the hours into three quantiles, using the middle third of the distribution as the reference category. For the binary outcomes of any cases and outbreak, we used logistic regressions. For the count outcome of number of deaths, after comparing fit of several possible count data options, we used hurdle negative binomial-2 regressions.¹¹ In addition to the staffing variables of primary interest, we controlled for facility size, ownership type, chain status, percentage of Medicaid residents, percentage of White residents, metropolitan status, and county cases per capita. Standard errors are clustered by state.

RESULTS

Unadjusted Associations

A total of 9,352 nursing facilities, or 71.0% of the sample, reported at least one COVID-19 case. Among facilities with cases, 27.5% experienced an outbreak, and the mean (median) number of deaths was 3.2 (0). We report the unadjusted associations between facility and county characteristics and COVID-19 cases/outbreaks in Table 1. Larger facilities, nonprofit ownership, fewer White residents, metropolitan county, and more county-level cases are associated with a higher probability of having any cases. The associations are of the same direction for probability of an outbreak for all characteristics except ownership: for-profit

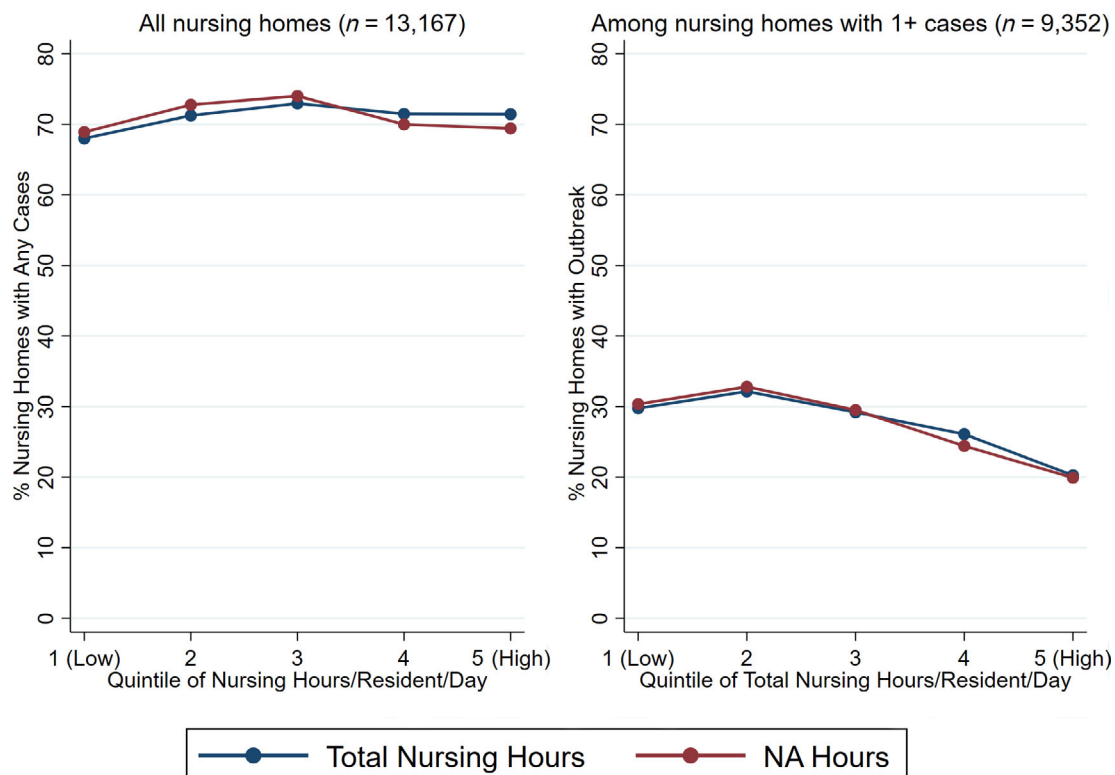


Figure 1. Unadjusted percentage of nursing homes with any cases and outbreak by quintile of nurse aide (NA) and total nursing hours. Definition of outbreak: >10% confirmed cases/beds or >20% confirmed plus suspected cases/beds, or ≥10 deaths. Quintiles were calculated using case-mix adjusted hours.

Table 2. Multivariable Regression Results: Effects of Staffing Measures on Outcomes

	Any cases Odds ratios		Outbreak Odds ratios		No. of deaths Marginal effects	
	(1)	(2)	(3)	(4)	(5)	(6)
Low NA hours	.887 (.058)		1.001 (.078)		-.034 (.184)	
High NA hours	1.027 (.071)		.790 (.058) ^a		-.981 (.229) ^a	
Low LPN hours	.975 (.052)		.847 (.073)		-.702 (.203) ^a	
High LPN hours	1.083 (.066)		1.064 (.081)		-.183 (.197)	
Low RN hours	.838 (.069) ^b		.974 (.070)		-.415 (.196) ^b	
High RN hours	1.341 (.088) ^a		1.031 (.079)		-.243 (.217)	
Low total nursing hours		.827 (.071) ^b		.924 (.073)		-.371 (.186) ^b
High total nursing hours		1.153 (.109)		.822 (.057) ^a		-1.059 (.229) ^a
Low RN/Total nursing hours		.887 (.052) ^b		1.018 (.062)		-.389 (.207)
High RN/Total Nursing hours		1.218 (.078) ^a		1.034 (.069)		-.296 (.195)
County cases/residents, Q2	1.344 (.117) ^a	1.345 (.117) ^a	1.210 (.168)	1.230 (.173)	.076 (.209)	.059 (.208)
County cases/residents, Q3	1.838 (.169) ^a	1.858 (.169) ^a	2.059 (.254) ^a	2.138 (.270) ^a	1.206 (.233) ^a	1.226 (.236) ^a
County cases/residents, Q4	3.316 (.521) ^a	3.320 (.518) ^a	3.890 (.518) ^a	4.013 (.510) ^a	2.687 (.252) ^a	2.719 (.254) ^a
County cases/residents, Q5 (high)	6.190 (1.085) ^a	6.204 (1.079) ^a	6.171 (1.052) ^a	6.318 (1.114) ^a	5.033 (.274) ^a	5.096 (.279) ^a
N	12,117	12,117	8,626	8,626	8, 626	8, 626

Note: Definition of outbreak: >10% confirmed cases/beds or >20% confirmed plus suspected cases/beds, or ≥10 deaths.

Standard errors clustered by state.

Regression also controls for nursing facility and county characteristics (omitted here for brevity).

Columns 1 and 2 include all Nursing Home Compare facilities with COVID-19 reporting and complete staffing rating information.

Columns 3, 4, 5, and 6 are limited to facilities with at least one case.

Low [High] hours = less [greater] than 33rd [66th] percentile of case-mix adjusted hours.

^aP<.01.

^bP<.05.

ownership is associated with a higher probability of an outbreak among facilities with at least one case. Not being part of a chain and a higher share of Medicaid residents are associated with higher probability of an outbreak. Facilities report whether there is access to COVID-19 testing for residents and PPE and staffing shortages. Access to testing is higher and shortages are more prevalent among facilities with cases and outbreaks. Overall NHC star ratings are similar between facilities with cases and outbreaks and those without. Facilities without cases/outbreaks have marginally higher health inspection ratings and lower quality measure ratings. Interestingly, staffing ratings are higher among facilities with cases than without, but among facilities with cases, staffing ratings and NA, total nurse and RN/total nurse hours are higher among facilities without outbreaks than facilities with outbreaks.

We explore this unadjusted relationship between staffing and cases/outbreaks graphically in Figure 1. The percentage of facilities with any cases is approximately 70% for all NA hours quintiles. However, among facilities with at least one case, 31.6% of facilities in the lower 40th percentile of the distribution of NA hours have experienced an outbreak compared with only 22.1% in the upper 40th percentile. Similarly, the mean number of deaths is 4.1 in quintiles 1 and 2 and 2.0 in quintiles 4 and 5 (results not shown). The pattern in which more staffing is not associated with the probability of any case but is associated with fewer outbreaks is also evident for total nursing hours. We replicate this figure for the other measures of staffing in Supplementary Figures S1-S3. LPN, RN, and RN/total

nursing hours do not exhibit these same associations with cases and outbreaks.

Regression-Adjusted Associations

Regression-adjusted results are presented in Table 2. High RN hours are associated with higher probability of a facility experiencing any cases. Among facilities with at least one case, high NA hours and high total nursing hours are associated with a lower probability of outbreak and fewer deaths. Low LPN hours is associated with fewer deaths but not the probability of an outbreak. RN hours and RN/total nursing hours are not associated with outbreak severity. We report the odds ratios/marginal effects for the facility and county characteristics in Supplementary Table S1. By far, the largest association with COVID-19 outcomes is the county-level number of cases in the community, that is, where the virus is circulating.

We report the results of several robustness checks in the Supplementary Appendix. First, we controlled for concurrent staffing and PPE shortages and resident access to testing, and associations between the outcomes and baseline staffing remain unchanged (Supplementary Table S2). Second, we repeated the analyses controlling for state. We found that both between-state and within-state variation in staffing contribute to the overall findings (Supplementary Table S4). Finally, the main analyses examined relationships between baseline staffing and cumulative cases and deaths. We used weekly cases and deaths for the week ending June 14 to explore if these relationships change over time; they did not (Supplementary Table S5).

DISCUSSION

We conducted the first national examination of the association of nursing home staffing levels and COVID-19 cases and outbreak severity using the newly available CMS COVID-19 Nursing Home dataset. We find a nuanced story: Staffing levels may play one role (or no role) in the probability that COVID-19 gets into a nursing home and another role in trying to stem transmission and deaths once it is in. For example, whereas staffing levels are unrelated to the probability of having at least one case or even increase that probability, perhaps by increasing traffic in and out of the facility, higher NA and total nursing hours are associated with a lower probability of a larger outbreak and fewer deaths.

The effect sizes of staffing are relatively small. For example, being in the top third of the distribution of NA hours is associated with one fewer death; similarly, being in the top third of total nursing hours is associated with 1.1 fewer deaths. Although potentially meaningful, these effects are dwarfed by the effect of where the virus is circulating. The largest magnitude effects we found are for county metropolitan status and county-level number of COVID-19 cases per capita that occur among the general population. Compared with counties in the lowest 20th percentile of per capita cases, counties in the highest 20th percentile experience 6.2 times the odds of detecting any cases, and among facilities with one or more cases, 6.2 times the odds of an outbreak or five additional deaths.

It is important to acknowledge the limitations of the COVID-19 Nursing Home data. CMS required reporting beginning May 8, and facilities have the option to report cases/deaths going back to January 1. As a result, the CMS data on total cases/deaths represent an undercount, especially in states that experienced early outbreaks.

These limitations aside, using the newer national CMS COVID-19 data, we confirm the conclusions of earlier, more geographically limited studies²⁻⁴ that found higher NHC ratings are not associated with better COVID-19 outcomes (Supplementary S3). The prevalence of the virus in the community remains the strongest predictor of how nursing homes fare under the pandemic, and this fact should guide policies about assistance and resource distribution.

At the same time, although staffing ratios are part of the star ratings, our detailed examination found a nuanced story of the role of nursing home staff. As the pandemic continues and best practices emerge, the role of staffing is likely to take on increased importance. Implementing efforts to stem transmission, such as regular testing and cohorting of both residents and staff, is difficult without sufficient staffing levels. With states relaxing visitor restrictions, the daily traffic of staff in and out of the facility, and asymptomatic spread, keeping COVID-19 out of nursing homes may be close to impossible. However, along with sufficient PPE and regular testing, having enough NAs to implement virus containment will be crucial if deaths are to be averted. Although nursing homes in neighborhoods with the highest virus prevalence will continue to be the most vulnerable and

in need of assistance in the short run, a critical and longer run focus must be to improve nursing home staffing. This is a long-standing challenge that is likely to become even more important as the pandemic continues.

ACKNOWLEDGMENTS

Financial Disclosure: Rebecca Gorges was supported by the Agency for Healthcare Research and Quality under Grant Award No. T32 HS000084. The contents are solely the responsibility of the authors and do not necessarily represent the official views of the Agency for Healthcare Research and Quality.

Conflict of Interest: The authors have declared no conflicts of interest for this article.

Author Contributions: All the authors contributed to develop the study concept and design. Rebecca Gorges performed the data analysis. All authors contributed to the interpretation of the data and preparation of the manuscript.

Sponsor's Role: Not applicable.

REFERENCES

- 43% of U.S. coronavirus deaths are linked to nursing homes. *New York Times*. <https://www.nytimes.com/interactive/2020/us/coronavirus-nursing-homes.html>. Accessed July 1, 2020.
- Konetzka RT. Testimony of R. Tamara Konetzka, PhD. 2020. https://www.aging.senate.gov/imo/media/doc/SCA_Konetzka_05_21_20.pdf. Accessed July 4, 2020.
- Abrams HR, Loomer L, Gandhi A, Grabowski DC. Characteristics of U.S. nursing homes with COVID-19 cases. *J Am Geriatr Soc*. 2020;68:1653–1656. <https://doi.org/10.1111/jgs.16661>
- White EM, Kosar CM, Feifer RA, et al. Variation in SARS-CoV-2 prevalence in US skilled nursing facilities. *J Am Geriatr Soc*. 2020. <https://doi.org/10.1111/jgs.16662>
- Harrington C, Ross L, Chapman S, Halifax E, Spurlock B, Bakerjian D. Nurse staffing and coronavirus infections in California nursing homes. *Policy Polit Nurs Pract*. 2020;21:174–186. <https://doi.org/10.1177/1527154420938707>
- Li Y, Temkin-Greener H, Gao S, Cai X. COVID-19 infections and deaths among Connecticut nursing home residents: facility correlates. *J Am Geriatr Soc*. 2020;68(9). <https://doi.org/10.1111/jgs.16689>
- COVID-19 Nursing Home Data. <https://data.cms.gov/stories/s/COVID-19-Nursing-Home-Data/bkwz-xpvg/>. Accessed July 1, 2020.
- Shaping Long Term Care in America Project at Brown University funded in part by the National Institute on Aging (1P01AG027296). <http://ltcfocus.org/>. Accessed May 4, 2020.
- Coronavirus locations: COVID-19 map by county and state. USAFacts.org. <https://usafacts.org/visualizations/coronavirus-covid-19-spread-map>. Accessed June 30, 2020.
- Center for Clinical Standards and Quality/Quality, Safety & Oversight Group. COVID-19 survey activities, CARES act funding, enhanced enforcement for infection control deficiencies, and quality improvement activities in nursing homes. 2020. <https://www.cms.gov/files/document/qso-20-31-all.pdf>. Accessed June 7, 2020.
- Deb P, Norton EC, Manning WG. *Health Econometrics Using Stata*. College Station, TX: Stata Press; 2017.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Supplementary Appendix S1: Supporting Information.