Incidence and prevalence of COVID-19 within a healthcare worker cohort during the first year of the SARS-CoV-2 pandemic

Sarah B. Doernberg, MD,[1*] Marisa Holubar, MD,[2*] Vivek Jain, MD,[3*] Yingjie Weng, MHS,[4] Di Lu, MS,[4] Jenna B. Bollyky, MD,[5] Hannah Sample, BS,[6] Beatrice Huang, BA,[7] Charles S. Craik, PhD,[8] Manisha Desai, PhD,[4] George W. Rutherford, MD,[9†] Yvonne Maldonado, MD[5†] and the CHART Study Consortium[**]

[1]Division of Infectious Diseases, University of California, San Francisco (UCSF); [2]Division of Infectious Diseases and Geographic Medicine, Stanford University School of Medicine, [3]Division of HIV, Infectious Diseases & Global Medicine, San Francisco General Hospital, University of California, San Francisco (UCSF); [4]Quantitative Sciences Unit, Stanford University School of Medicine; [5]Division of Pediatric Infectious Diseases, Stanford University School of Medicine; [6]Department of Biochemistry and Biophysics (UCSF); [7]Department of Family and Community Medicine, San Francisco General Hospital, University of California, San Francisco (UCSF); [8]Department of Pharmaceutical Chemistry (UCSF), [9]Division of Infectious Disease and Global Epidemiology, Department of Epidemiology and Biostatistics, UCSF

[*†] Authors contributed equally to this report

[**] Consortium collaborators listed in acknowledgment section

Corresponding author:

Marisa Holubar, MD MS

Clinical Associate Professor of Medicine, Division of Infectious Diseases and Geographic Medicine

Stanford University School of Medicine 300 Pasteur Drive, Lane Building 145 Stanford, CA 94305

mholubar@stanford.edu

Summary

In a longitudinal cohort study of 2435 Bay Area healthcare workers conducted from July 2020-January 2021, COVID-19 incidence was low. COVID-19 was strongly associated with community COVID-19 contacts but was not associated with work contacts unless accompanied by high-risk exposure.

Abstract

Background: Preventing SARS-CoV2 infections in healthcare workers (HCWs) is critical for healthcare delivery. We aimed to estimate and characterize the prevalence and incidence of COVID-19 in a US HCW cohort and to identify risk factors associated with infection.

Methods: We conducted a longitudinal cohort study of HCWs at 3 Bay Area medical centers using serial surveys and SARS-CoV-2 viral and orthogonal serological testing, including measurement of neutralizing antibodies. We estimated baseline prevalence and cumulative incidence of COVID-19. We performed multivariable Cox proportional hazards models to estimate associations of baseline factors with incident infections and evaluated the impact of time-varying exposures on time to COVID-19 using marginal structural models.

Results: 2435 HCWs contributed 768 person years of follow-up time. We identified 21/2435 individuals with prevalent infection, resulting in a baseline prevalence of 0.86% (95% CI, 0.53% to 1.32%). We identified 70/2414 (2.9%) incident infections yielding a cumulative incidence rate of 9.11 cases per 100 person years (95% CI 7.11 to 11.52). Community contact with a known COVID-19 case most strongly correlated with increased hazard for infection (HR 8.1, 95% CI, 3.8, 17.5). High-risk work-related exposures (i.e., breach in protective measures) drove an association between work exposure and infection (HR 2.5, 95% CI, 1.3-4.8). More cases were identified in HCW when community case rates were high.

Conclusion: We observed modest COVID-19 incidence despite consistent exposure at work. Community contact was strongly associated with infections but contact at work was not unless accompanied by high-risk exposure.

Keywords: COVID-19, SARS-CoV2, healthcare worker, healthcare personnel

Introduction

Many assume that healthcare workers (HCWs) acquire COVID-19 at work.[1–3] While early studies supported work-related risks, more recent studies have shown that factors including community-based exposures, race/ethnicity, and residential zip code, are associated with SARS-CoV2 acquisition and may be more consequential than workplace exposures.[4–6]

Existing literature and media reports have noted widely varying estimates of prevalence, incidence, and risk factors for infection.[7–9] Few studies have assessed HCW infection and risk longitudinally despite changes in community prevalence of COVID-19 over time, workplace infection prevention efforts, and dynamic individual adherence to public health measures outside of work. Two European groups reported data from longitudinal HCW screening programs and estimated prevalence and incidence of infection[2,10] but to date, a similar granular approach to describing prevalence and incidence of COVID-19 in United States HCWs has not been reported.

Additionally, most seroprevalence studies of HCWs—both cross-sectional and longitudinal—have used a single unconfirmed serology test without orthogonal confirmation (i.e., using a different test) as their main outcome measure. Most studies have also not reported neutralizing antibody titers.[11–14] Orthogonal antibody testing increases specificity, which is critical when testing populations with low disease prevalence.[15] Further, neutralizing antibody titers provides a functional assessment of immune responses.

To address these current gaps in our understanding of SARS-CoV-2, we sought to estimate and characterize the prevalence and incidence of COVID-19 using both reverse-transcription polymerase

chain reaction (RT-PCR) and orthogonal antibody testing in a large longitudinal cohort of HCW during a dynamic phase of the US epidemic and to identify risk factors associated with HCW infection.

Methods

Ethics statement

The COVID-19 Healthcare Worker Antibody and RT-PCR Tracking (CHART) Study was approved by the University of California, San Francisco's Committee on Human Subjects Research, and the Stanford University School of Medicine Panel on Human Subjects in Medical Research.

Study population and setting

From May-September 2020, we recruited HCW from Stanford Health (SHC), UCSF Health (UCSF), and Zuckerberg San Francisco General Hospital (ZSFGH) for this longitudinal, prospective cohort study. These three medical centers serve large, mostly non-overlapping catchment populations in the San Francisco Bay Area and implemented similar mitigation policies over time (Supplementary Table 1). Recruitment included medical center-wide email and verbal announcements, targeted email notifications to department leaders, and recruitment flyers.

HCWs completed an electronic screening questionnaire (Supplementary Material). Inclusion criteria were (1) age ≥18 years old, (2) employment at one of the three medical centers, and (3) did not anticipate ending employment or taking leave in the next 6 months. Eligible HCWs provided consent electronically. We collected study data using REDcap electronic data capture tools hosted at Stanford University.[16,17]

Schedule of evaluations

The study was conducted from July 2020 to January 2021. Participants completed up to 10 visits: 7 visits at two-week intervals (±7 days) followed by 3 visits at four-week intervals (±7 days) up to completion or end of the study.. At all visits, participants completed an electronic survey and study staff collected nasopharyngeal (NP) swabs; swabs were optional for the final 3 visits. Participants underwent phlebotomy monthly for anti-SARS-CoV-2 antibody testing. For individuals who tested positive by either RT-PCR or serology, additional visits were scheduled weekly for four visits for serology testing only. Participants received no incentives or compensation for joining the study.

Laboratory RT-PCR and serology testing

The UCSF Clinical Laboratories and Chan-Zuckerberg Biohub analyzed samples from the UCSF and ZSFGH sub-cohorts. Serology was performed using an assay to detect anti-nucleocapsid IgG (anti-nucleocapsid Ab; Abbott Architect, Abbott Laboratories, Abbott Park, IL)[18]. The Stanford Clinical Virology Laboratory analyzed samples using an assay to detect anti-spike IgG (anti-spike Ab; Eurimmune Medizinische Labordiagnostika AG, Lübeck, Germany)[19] as well as a laboratory-derived assay to detect anti-receptor-binding-domain IgG (anti-RBD Ab)[20]. Samples that were positive at one laboratory underwent confirmatory testing at the other laboratory. Serum samples that were positive for antibodies to either spike, nucleocapsid, or both proteins were assayed for the presence of neutralizing antibodies at UCSF or at Vitalant Research Institute (San Francisco, CA) by optimizing a lentivirus-based pseudotype neutralization assay[21].

At UCSF, RT-PCR testing was performed using either (1) the M2000 Abbott RealTime Sars-CoV2 assay[22] amplifying RdRP and N genes, (2) the MAGPIX Luminex NxTag CoV Extended Panel assay (Austin, TX)[23] amplifying the N gene, the Orf1ab gene, and the E gene, or (3) a Clinical Laboratory Improvement Amendments (CLIA)-validated laboratory-derived test modified from the CDC amplifying the N and E genes.[24,25] At SHC, RT-PCR testing was performed using an SHC laboratory-derived test amplifying the E gene or the Panther Fusion SARS-CoV-2 assay (Hologic, Massachusetts).[26,27]

Definitions of COVID-19 exposures, positive test results, and cases

We defined a low-risk work exposure as providing direct care to, being within 6 feet of, directly interacting with the environment in which a COVID-19 patient received care, or processing laboratory samples from a COVID-19 patient. We defined a high-risk exposure at work as ever interacting with a COVID-19 patient without full PPE—the institutionally recommended PPE for care of patients with COVID-19—or having a breach in PPE (e.g., tears, accidental removal).

We defined an RT-PCR result as positive if the result was (1) detected or (2) indeterminate (positive RT-PCR followed by negative subsequent confirmatory RT-PCR test(s) done according to medical center occupational health protocols).

We defined positive confirmed serology as having an initial positive serology (anti-nucleocapsid Ab or anti-spike Ab) followed by confirmation with a second positive serology using a different target (antnucleocapsid Ab, anti-spike Ab, or neutralizing Ab). Positive confirmed serology represented prior COVID-19 infection. We defined a positive unconfirmed serology as an isolated positive anti-

nucleocapsid or anti-spike Ab (i.e., a negative result on confirmatory testing) in the absence of RT-PCR positivity.

We defined baseline prevalent cases as participants with positive RT-PCR or positive confirmed serology at their initial visit. Participants who did not have baseline infection entered an incident cohort. We defined incident cases among this cohort as participants with a positive RT-PCR or a positive confirmed serology at any subsequent visit. The date of incident infection was the first date on which either the RT-PCR or the first serology test was positive (if confirmatory testing occurred within 4 weeks).

Statistical analyses

We estimated the prevalence as the proportion of cases at baseline out of total number of enrolled participants who completed baseline visits. We estimated the cumulative incidence as the number of incident cases divided by the total follow-up time per 100 person-years and assumed a uniform incidence distribution across the 6-month follow-up time. We censored person-time when a participant met the case definition, completed or withdrew from the study, or received a first dose of any COVID-19 vaccine. We calculated the confidence intervals (CI) using a non-parametric bootstrapping method. We conducted a sensitivity analysis to assess the impact of different case definitions on estimates considering 1) all unconfirmed positive serology results as cases, 2) all individuals with a single positive RT-PCR, no positive serology result, and at least one serology measurement ≥4 weeks after the positive RT-PCR as potential false positives and removing them from case counts. We obtained community-wide data on COVID-19 incidence in the six Bay Area counties from the California Department of Public Health.[28]

We compared characteristics of prevalent and incident cases to non-cases. For binary time-varying exposures, we used participant self-report at the most recent visit prior to censoring. For continuous time-varying exposures, we computed median responses across all visits prior to censoring. We reported symptoms using the most recent reported status at the visit at which infection was identified. We reported standardized mean difference (SMD) to describe magnitude of difference in characteristics between incident cases and non-cases. Magnitude of effect is considered small if SMD=0.2, medium if SMD=0.5, and large if SMD=0.8.

In the incident cohort, we first assessed associations between time to infection and baseline characteristics using multivariable Cox proportional hazards models. We evaluated the impact of pre-specified time-varying exposures on time to infection via marginal structural models (MSM).[29–32] We implemented a 2-step MSM model for each time-varying exposure by first estimating inverse probability of treatment weights (IPTW), in which exposure probability was estimated for each participant at each visit conditioning on fixed and other time-varying exposures up to that time. To stabilize weights, we excluded correlated time-varying variables. Each participant was weighted with the inverse predicted probability of exposure to simulate a counterfactual participant. Second, we applied an extended Cox proportional hazard model with IPTWs and reported hazard ratios for the impact of time-varying exposures on time to infection. For all regression analyses, we imputed missing laboratory data using a last observation carrying forward method and missing time-invariant or time-varying data using multiple imputation. We controlled family wise type I error at 0.05 and used the significance level of 0.05 in hypothesis tests. All analysis were conducted using SAS 9.4.3 (SAS Institute, Research Triangle Park, NC) and R, version 4.5.3. (R Project for Statistical Computing, Vienna, Austria) by YW, DL, and ND.

Results

Healthcare Worker Demographics

Of 3918 individuals screened, 2435 provided consent and completed the first study visit, contributing 768 total person-years of follow-up time (Figure 1). Baseline demographics are presented in Table 1. Overall, mean age was 40.4 years (standard deviation (SD), 10.1), 1923/2435 (79%) were female, and most participants (1921/2435, 79%) reported providing direct patient care, including 701/1921 (36%) who performed AGPs. Many participants reported work-related COVID-19 exposure (1477/2419, 61%) with 797/1477 (54%) reporting high-risk exposure. Only 176/2419 (7%) participants overall reported contact with a COVID-19 positive person outside of work.

Overall, demographic and behavior characteristics of participants with prevalent and incident COVID-19 reflected overall cohort characteristics (Table 1), including 73/91 (80%) providing direct patient care, mostly as nurses (42/91, 46%) or clinicians (24/91, 26%). During the course of the study, HCW time spent in the healthcare environment and work-related exposures to COVID-19 were both stable (Figures 2A and 2B).

Prevalence and incidence of COVID-19

We identified 21/2435 individuals with evidence of COVID-19 at baseline and estimated a prevalence of 0.86% (95% CI, 0.53% to 1.32%). We identified 70/2414 (2.9%) individuals with incident COVID-19 during follow-up and estimated a cumulative incidence rate of 9.11 cases per 100 person years (95% CI 7.11 to 11.52). The number of incident cases increased with rising prevalence of COVID-19 in the 8-county region in which the study was conducted (Figure 2). Incidence rate estimates did not differ by sub-groups of gender, race/ethnicity, or job role (Supplementary Figure 1).

All 21 prevalent COVID-19 cases met the case definition with a positive serology; only 3 also had a positive RT-PCR. Most of the 70 incident cases were identified by a positive RT-PCR (53/70, 76%) with or without a positive serology. Of the 17/70 (24%) participants meeting the case definition by positive serology alone, only 2 (12%) individuals had a positive RT-PCR at a later visit (2 and 5 weeks after positive serology).

We performed a sensitivity analysis using an alternative incident COVID-19 case definition that included all unconfirmed positive serology results as cases, resulting in 26 prevalent and 71 incident cases. This slightly increased the baseline prevalence to 1.07% (95% CI, 0.79% to 1.56%) and increased the cumulative incidence rate to 9.26 cases per 100 person-years (95% CI, 7.24 to 11.69).

To examine the impact of potential false positive RT-PCRs, we performed a second sensitivity analysis using a second alternative case definition that excluded 7 cases meeting this definition. This decreased the cumulative incidence rate to 8.18 cases per 100 person-years (95% CI, 6.29 to 10.4).

Overall, the testing yield of the incident cohort was relatively low: only 30/12,007 (0.25%) RT-PCR tests performed on asymptomatic participants were positive, and 7/30 (23%) of these met the false positive case definition.

Figure 3 demonstrates the participant-level temporal sequence of testing results for all baseline prevalent cases and all incident cases. We found substantial evolution of antibody responses over time: of the 56 cases initially diagnosed by RT-PCR, 11 had at least one positive antibody at

diagnosis. By the end of follow-up, this rose to 27 individuals with at least one positive antibody test. Further, eleven individuals who had detected antibodies on one assay subsequently tested positive using another assay.

COVID-19 symptomology

Of the 70 incident cases, 36 (51.4%) were asymptomatic at diagnosis (30 with positive RT-PCR, 6with positive serology only.) Of the 36 participants who were asymptomatic at diagnosis, 14/22 (64%) who completed a follow-up symptom assessment had remained asymptomatic. Of the 25 participants who were symptomatic at diagnosis, 19 had positive RT-PCR and 6 had positive serology only.

Among the 1170 participants who reported symptoms at any visit, 58 (5%) were confirmed as prevalent or incident cases. Among 1252 participants who never reported symptoms, 32 (3%) were confirmed as prevalent or incident cases.

While incident cases more commonly reported ever having symptoms (48/70, 69%), many non-cases (1112/2344, 48%) reported symptoms at least once (Supplementary Table 2). The most common symptoms reported by non-cases were fatigue (326, 14%), headache (466, 20%), nasal congestion (325, 14%), and rhinorrhea (412, 17%). Non-cases infrequently reported fever, chills, or decreased taste/smell while cases reported them more commonly.

Predictors of COVID-19 infection

In a multivariable Cox proportional hazards model, we did not find an association of incident COVID-19 with fixed variables including baseline age, gender, race, ethnicity, household size, role, or work

category (Supplementary Table 3). In marginal structural models of self-reported time-varying variables, community contact with a known COVID-19 case strongly correlated with increased hazard for COVID-19 (HR 8.1, 95% CI, 3.8, 17.5; Table 2). Self-reported exposure to a COVID-19 patient at work was associated with infection (p=0.013), but this appeared to be primarily driven by high-risk exposures (i.e., a PPE failure or breach or an exposure to patient biological material; HR 2.5, 95% CI, 1.3-4.8). Increasing community COVID-19 case rate showed a trend towards elevated adjusted hazard of HCW infection, but this finding did not reach statistical significance (HR 1.3, 95% CI 0.97-1.8). Neither time spent in the healthcare workplace, time spent providing direct patient-facing care, or adherence to community mitigation strategies were associated with COVID-19 infection.

Discussion

In this large observational cohort of healthcare workers, we observed modest COVID-19 infection rates despite consistent COVID-19 exposure at work. Changes in COVID-19 incidence tracked most closely with community infection rates and self-reported community contact with known COVID-19 cases rather than work-related factors, except when breaches in standard safety protocols or PPE occurred.[5] Our data provide evidence of the overall safety of standard healthcare work environment protocols and PPE guidelines, and are concordant with emerging literature showing that the main COVID-19 related risks to HCW are those coming from home and community-based factors.[33]

By combining longitudinal and orthogonal RT-PCR and serology testing, our study allowed for a robust granular estimation of the true incidence of COVID-19 infection among HCW. Unlike many studies based on a single serologic test, we used confirmatory serology testing and also measured neutralizing antibody responses.[34] As our data show, the serological response to infection is

multifaceted and evolves over time; measuring a single antibody response to one target may result in inaccurate estimates of true infection rates.[35] By testing serially and confirming antibody responses, we captured some COVID-19 cases that would have likely been missed with a single test in time and excluded others that were likely false positives.

In our sensitivity analysis accounting for potential false positive COVID-19 cases, our incidence estimates were 10% lower. This may have been an underestimate because many cases were diagnosed at the end of the study and lacked follow-up time to differentiate true from false positives. Misclassifying false positive test results as true cases can impact the ability of a healthcare system to operate by limiting critical staffing and can also have adverse implications for household contacts of HCW. COVID-19 screening programs for HCW must balance the value of prompt diagnosis with the downside of potential false positive results.

Our study is subject to several limitations. We enrolled volunteer participants and had a high fraction of MD, MD-equivalent, and RN practitioners. This cohort composition did not comprehensively reflect the occupational diversity within our medical centers. Thirty-eight percent of those screened did not enroll in the study; because we did not assess reasons for non-participation, it is unclear to what degree this may have introduced any bias in our study population. We relied on self-reporting of COVID-19 related risks both at work and home, which may have resulted in overreporting of adherence to protective measures. Additionally, our institutional PPE recommendations changed over time; as such, not all breaches in PPE are considered equivalent. However, unlike many studies that have used information from employee health and safety offices, our study was independent of the medical centers in order to foster confidential no-fault reporting. Because sequencing of virus was beyond the scope of the study, the association between self-

reported breach in PPE and incident COVID-19 cases remains solely an association and not proof that the breach itself led to the incident infection. Additionally, we did not perform orthogonal SARS-CoV-2 antibody testing on samples that were initially antibody negative, and thus could have missed certain incident cases. We also did not include confirmatory testing of RT-PCR results so could have inadvertently included false-positive RT-PCR results in incidence rate estimates. We addressed this with a sensitivity analysis and found that incidence rates were minimally impacted. Finally, our study was conducted before more recent variants of concern with increased transmissibility and immune escape emerged. One key strength of our study was our use of marginal structural modeling using detailed longitudinal data to better estimate risks.

Within a large group of frontline healthcare workers, our data indicate that healthcare workplaces pursuing comprehensive mitigation strategies can operate safely despite facing sequential waves of COVID-19 cases. However, HCW do face community-based risks for acquiring COVID-19. Medical center infection control practices, vaccination programs, and community mitigation approaches should be sustained and maximized to protect HCW and health systems during periods of future risk related to rising caseloads and emerging SARS-CoV-2 variants.

NOTES

Acknowledgements

We sincerely thank all participants for generously volunteering their time and effort to this study. We also thank all study staff, our participating health systems including occupational health and infection control and prevention teams, medical center staff and university and medical center leadership.

Authors also acknowledge Salu Ribiero and Bay Area Phlebotomy Group, UCSF colleagues Tim Moriarty, Chantal Giunta, Erik Stewart, Robert Cusack, Alexander Lee, Miguel Ortiz, Peter Suslow, Don Rose, Pilar Collins, Eunice Stephens, Tatiana Urisman, Dan Henroid, Pamela Hudson, Jennifer Dearman, Leigh Ann Ambrose, Tim Judson, Sabrina Mann; ZSFG colleagues, Geary Wong, Jeff Whitman, Barbara Haller, Kara Lynch, Jeff Schmidt, Lisa Winston, Elaine Dekker, Mark Jacobson; and Stanford Healthcare colleagues Tim Morrison, and Karen Baines.

The Stanford REDCap platform (http://redcap.stanford.edu) is developed and operated by Stanford Medicine Research IT team. The REDCap platform services at Stanford are subsidized by a) Stanford School of Medicine Research Office, and b) the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through grant UL1 TR001085†. REDCap (Research Electronic Data Capture) is a secure, web-based software platform designed to support data capture for research studies, providing 1) an intuitive interface for validated data capture; 2) audit trails for tracking data manipulation and export procedures; 3) automated export procedures for seamless data downloads to common statistical packages; and 4) procedures for data integration and interoperability with external sources

Authors acknowledge the Chan Zuckerberg Initiative for their support.

List of CHART Study Consortium collaborators

University of California, San Francisco (UCSF): Parul Bhargava, Markus Bohn, Jessica Chao, Charles Craik, Sarah B. Doernberg, Jacob Ghahremani, David Glidden, Ralph Gonzales, Beatrice Huang, Sravya Jaladanki, Aida Julien, Daniel Lowenstein, Steve Miller, Audrey Mustoe, Marcus Paoletti, George W. Rutherford, Hannah Sample, Rodolfo Villa, Emerald Wan, Aimee Williams

San Francisco General Hospital, UCSF: Lillian Brown, Jessica Chuang, Vivek Jain, Carina Marquez, Guntas Padda, Luis Rubio, Daisy Valdivieso

Stanford Medicine/Stanford Health Care: Rosebay Abad, Anthony Bet, Jenna Bollyky, Manisha Desai, Jeffrey Fung, Anna Graber, Cole Holderman, Marisa Holubar, Hannah Kelley, Amanda Kempema, Christina Kong, Christopher Leung, Joseph Lohmann, Di Lu, Yvonne Maldonado, Lloyd Minor, Lorena Orozco, Benjamin A. Pinsky, Jamie Saxeena, Matthew Sklar, Hilary Tang, Jasmine Wiese, Yingjie Weng

Chan-Zuckerberg Biohub: Emily Crawford, Joe DeRisi, All members of the CLIAHUB Consortium

Funding

Funding provided by the Chan Zuckerberg Initiative Foundation (CZIF2020-003870).

Conflicts of Interest

Sarah B. Doernberg, MD: S.B.D. has received COVID-research funding from Gilead and NIH and has served as a consultant to and has received non-COVID research funding from Genentech, Basilea, and NIH

George W. Rutherford, MD: G.R has received grant funding from Centers for Disease Control and Prevention, California Department of Public Health, has served as a consult regarding adverse events

for Moderna, and has provided expert declarations and depositions for the California Department of Justice.

Yvonne Maldonado, MD: Y.M. has received grant funding from Pfizer (C3671008; C4591007), NIH (U54 MD010724; U54 MD010724- 05S1; R21AI148810; P30AG059307; 000522211-022), Bill & Melinda Gates Foundation (OPP1113682), and the Chan Zuckerberg Foundation (12089sc), and has served as an advisor for and received payment or honoraria from the American Academy of Pediatrics, has served on a data safety monitory board for Pfizer, and Data Safety Monitoring Board.

No conflicts for Marisa Holubar, MD; Vivek Jain, MD; Yingjie Weng, MHS; Di Lu, MS; Jenna B. Bollyky, MD; Hannah Sample, BS; Beatrice Huang, BA; Charles S. Craik, PhD; and Manisha Desai, PhD

References

1. Self WH, Tenforde MW, Stubblefield WB, et al. Decline in SARS-CoV-2 Antibodies After Mild Infection Among Frontline Health Care Personnel in a Multistate Hospital Network — 12 States, April–August 2020. MMWR Morb Mortal Wkly Rep 2020;69(47):1762–6.

2. Weinberger T, Steffen J, Osterman A, et al. Prospective Longitudinal Serosurvey of Health Care Workers in the First Wave of the SARS-CoV-2 Pandemic in a Quaternary Care Hospital in Munich, Germany. Clin Infect Dis Off Publ Infect Dis Soc Am 2021;ciaa1935.

3. Moncunill G, Mayor A, Santano R, et al. SARS-CoV-2 Seroprevalence and Antibody Kinetics Among Health Care Workers in a Spanish Hospital After 3 Months of Follow-up. J Infect Dis 2021;223(1):62–71.

4. Baker JM, Nelson KN, Overton E, et al. Quantification of Occupational and Community Risk Factors for SARS-CoV-2 Seropositivity Among Health Care Workers in a Large U.S. Health Care System. Ann Intern Med [Internet] 2021 [cited 2021 Feb 21];Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7877798/

5. Fell A. SARS-CoV-2 Exposure and Infection Among Health Care Personnel — Minnesota, March 6–July 11, 2020. MMWR Morb Mortal Wkly Rep [Internet] 2020 [cited 2021 Aug 4];69. Available from: https://www.cdc.gov/mmwr/volumes/69/wr/mm6943a5.htm

6. Jacob JT, Baker JM, Fridkin SK, et al. Risk Factors Associated With SARS-CoV-2 Seropositivity Among US Health Care Personnel. JAMA Netw Open 2021;4(3):e211283–e211283.

7. Chou R, Dana T, Buckley DI, Selph S, Fu R, Totten AM. Update Alert 7: Epidemiology of and Risk Factors for Coronavirus Infection in Health Care Workers. Ann Intern Med [Internet] 2021 [cited 2021 Feb 21];Available from: https://www.acpjournals.org/doi/10.7326/L21-0034

8. Our key findings about US healthcare worker deaths in the pandemic's first year [Internet]. The Guardian. [cited 2021 Aug 2];Available from: https://www.theguardian.com/us-news/ng-interactive/2020/dec/22/lost-on-the-frontline-our-findings-to-date

9. CDC. COVID Data Tracker [Internet]. Cent. Dis. Control Prev. 2020 [cited 2021 Aug 2];Available from: https://covid.cdc.gov/covid-data-tracker

10. Lumley SF, O'Donnell D, Stoesser NE, et al. Antibody Status and Incidence of SARS-CoV-2 Infection in Health Care Workers. N Engl J Med 2021;384(6):533–40.

11. Herzberg J, Vollmer T, Fischer B, et al. Prospective Sero-epidemiological Evaluation of SARS-CoV-2 among Health Care Workers in a German Secondary Care Hospital. Int J Infect Dis 2021;102:136–43.

12. Jespersen S, Mikkelsen S, Greve T, et al. SARS-CoV-2 seroprevalence survey among 17,971 healthcare and administrative personnel at hospitals, pre-hospital services, and specialist practitioners in the Central Denmark Region. Clin Infect Dis Off Publ Infect Dis Soc Am 2020;ciaa1471.

13. Patel MM, Thornburg NJ, Stubblefield WB, et al. Change in Antibodies to SARS-CoV-2 Over 60 Days Among Health Care Personnel in Nashville, Tennessee. JAMA 2020;324(17):1781–2.

14. Martin C, Montesinos I, Dauby N, et al. Dynamics of SARS-CoV-2 RT-PCR positivity and seroprevalence among high-risk healthcare workers and hospital staff. J Hosp Infect 2020;106(1):102–6.

15. Xu G, Emanuel AJ, Nadig S, et al. Evaluation of Orthogonal Testing Algorithm for Detection of SARS-CoV-2 IgG Antibodies. Clin Chem 2020;66(12):1531–7.

16. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)--a metadata-driven methodology and workflow process for providing translational research informatics support. J Biomed Inform 2009;42(2):377–81.

17. Harris PA, Taylor R, Minor BL, et al. The REDCap consortium: Building an international community of software platform partners. J Biomed Inform 2019;95:103208.

18. Abbott. SARS-CoV-2 IgG Assay - for use with Architect [Internet]. [cited 2021 Aug 26];Available from: https://www.fda.gov/media/137383/download

19. FDA. Fact Sheet for Healthcare Providers - EUROIMMUN Anti-SARS-CoV-2 ELISA (IgG) [Internet]. [cited 2021 Aug 26];Available from: https://www.fda.gov/media/137607/download

20. Röltgen K, Powell AE, Wirz OF, et al. Defining the features and duration of antibody responses to SARS-CoV-2 infection associated with disease severity and outcome. Sci Immunol [Internet] 2020 [cited 2021 Aug 2];5(54). Available from: https://immunology.sciencemag.org/content/5/54/eabe0240

21. Ng DL, Goldgof GM, Shy BR, et al. SARS-CoV-2 seroprevalence and neutralizing activity in donor and patient blood. Nat Commun 2020;11(1):4698.

22. Abbott. Abbott RealTime SARS-CoV-2 - Emergency Use Authorization (EUA) [Internet]. [cited 2021 Aug 26];Available from: https://www.molecular.abbott/sal/9N77-095_SARS-CoV-2_US_EUA_Amp_PI.pdf

23. Luminex. NxTAG CoV Extended Panel Assay Package Insert [Internet]. [cited 2021 Aug 26];Available from: https://www.fda.gov/media/136500/download

24. Centers for Disease Control and Prevention Division of Viral Diseases. CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel [Internet]. 2021 [cited 2021 Sep 9];Available from: https://www.fda.gov/media/134922/download

25. Crawford ED, Acosta I, Ahyong V, et al. Rapid deployment of SARS-CoV-2 testing: The CLIAHUB. PLOS Pathog 2020;16(10):e1008966.

26. Stanford Health Care Clinical Virology Laboratory. Stanford Health Care Clinical Virology Laboratory SARS-CoV-2 test EUA Summary [Internet]. [cited 2021 Aug 26];Available from: https://www.fda.gov/media/136818/download

27. Hologic. SARS-CoV-2 Assay (Panther Fusion System) [Internet]. [cited 2021 Sep 15];Available from: https://www.hologic.com/sites/default/files/2020-03/AW-21159-001_002_01.pdf

28. Severe acute respiratory coronavirus virus 2 (SARS-CoV-2) seroprevalence in healthcare personnel in northern California early in the coronavirus disease 2019 (COVID-19) pandemic | Infection Control & Hospital Epidemiology | Cambridge Core [Internet]. [cited 2021 Aug 2];Available from: https://www.cambridge.org/core/journals/infection-control-and-hospital-epidemiology/article/severe-acute-respiratory-coronavirus-virus-2-sarscov2-seroprevalence-in-healthcare-personnel-in-northern-california-early-in-the-coronavirus-disease-2019-covid19-pandemic/6095A8BED6679CB8032561BDCF2C26F4

29. Robins JM, Hernán MÁ, Brumback B. Marginal Structural Models and Causal Inference in Epidemiology. Epidemiology 2000;11(5):550–60.

30. Hernán MA. A definition of causal effect for epidemiological research. J Epidemiol Community Health 2004;58(4):265–71.

31. Wal WM van der, Geskus RB. ipw: An R Package for Inverse Probability Weighting. J Stat Softw 2011;43(1):1–23.

32. Cole SR, Hernán MA. Constructing Inverse Probability Weights for Marginal Structural Models. Am J Epidemiol 2008;168(6):656–64.

33. Nguyen LH, Drew DA, Graham MS, et al. Risk of COVID-19 among front-line health-care workers and the general community: a prospective cohort study. Lancet Public Health 2020;5(9):e475–83.

34. Humoral Immune Response to SARS-CoV-2 in Iceland | NEJM [Internet]. [cited 2021 Aug 2];Available from: https://www.nejm.org/doi/10.1056/NEJMoa2026116?url_ver=Z39.88-2003&rfr_id=ori:rid:crossref.org&rfr_dat=cr_pub%20%200pubmed

35. Immune Responses in Patients with COVID-19: An Overview | Pediatric Annals [Internet]. [cited 2021 Aug 2];Available from: https://journals.healio.com/doi/abs/10.3928/19382359-20210421-02

Figure Legends

Figure 1. Participant flow diagram. Positive confirmed serology: initial positive serology (anti-nucleocapsid antibody (Ab) or anti-spike Ab) followed by confirmation with a second positive serology using a different target (anti-nucleocapsid Ab, anti-spike Ab, or neutralizing Ab). Positive unconfirmed serology: An isolated positive anti-nucleocapsid or anti-spike Ab (i.e., a negative result on confirmatory testing) in the absence of RT-PCR positivity. Prevalent COVID-19 cases: Participants with positive RT-PCR or a positive confirmed serology at baseline. Incident COVID-19 cases: Participants with a positive RT-PCR or a positive confirmed serology at any subsequent visit.

Figure 2. Work and community-related COVID-19 exposures and incident cases over time. Figure 2A-C. Self-reported work and home exposures over time. Each line depicts the 7-day smoothed median responses of each self-reported home or community behavior or exposure. The gray shading represents the 95% confidence interval around the average. **Fig. 2D. Incident cases in the context of surrounding community caseload.** Boxes indicate unique incident cases and are color coded by how they met case definition. The line depicts a 7-day smoothed average of community reported cases from the six San Francisco Bay Area Counties surrounding the three medical centers.

Figure 3. Timing and sequence of positive tests among healthcare workers with COVID-19. Each row represents all test results for each prevalent and incident case over the study period. Gray shading indicates each participant's follow-up time. Dots represent RT-PCR results and boxes represent serology results. Blue coloring indicates a negative PCR or serology; red coloring indicates a positive RT-PCR or confirmed positive serology. Orange boxes represent unconfirmed positive serology. Red box thickness correlates with the number of positive confirmed serologies (e.g., 2 or 3 positive antibody tests).

Table 1. Characteristics of healthcare workers

	Overall	Prevalent positive	Incident cohort		SMD*
			Incident negative	Incident positive	
n	**2435**	**21**	**2344**	**70**	
Follow-up time (person-years)	779	0	750	17.5	0.60
Age, years (mean [SD])	40.4 (10.1)	42.7 (13.0)	40.4 (10.05)	40.6 (11.73)	0.02
Health system					0.16
SHC	891 (36.6)	5 (23.8)	855 (36.5)	31 (44.3)	
UCSF Health	826 (33.9)	7 (33.3)	798 (34.0)	21 (30.0)	
ZSFGH	718 (29.5)	9 (42.9)	691 (29.5)	18 (25.7)	
Gender					0.11
Female	1923 (79.0)	13 (61.9)	1855 (79.1)	55 (78.6)	
Latinx Ethnicity					0.12
No	2036 (87.8)	17 (100.0)	1960 (87.7)	7 (10.6)	
Yes	267 (11.5)	0 (0.0)	260 (11.6)	7 (10.6)	
Decline to answer	15 (0.6)	0 (0.0)	15 (0.7)	0 (0.0)	
Race					0.14
White	1411 (60.9)	9 (52.9)	1364 (61.1)	38 (57.6)	
Asian	530 (22.9)	3 (17.6)	511 (22.9)	16 (24.2)	
Black	37 (1.6)	1 (5.9)	34 (1.5)	2 (3.0)	
Multiple races	137 (5.9)	3 (17.6)	131 (5.9)	3 (4.5)	
Other	143 (6.2)	0 (0.0)	138 (6.2)	5 (7.6)	
Decline to answer	58 (2.5)	1 (5.9)	55 (2.5)	2 (3.0)	
Highest level of education					0.29
Less than college	33 (1.4)	0 (0.0)	33 (1.5)	0 (0.0)	
College	1006 (43.4)	6 (35.3)	967 (43.3)	33 (50.0)	
Higher than college	1264 (54.5)	11 (64.7)	1222 (54.7)	31 (47.0)	
Other	15 (0.6)	0 (0.0)	13 (0.6)	2 (3.0)	
Co-morbid conditions					
None reported	1642 (71.5)	12 (70.6)	1582 (71.5)	48 (72.7)	0.03
Asthma/COPD	328 (14.3)	3 (17.6)	314 (14.2)	11 (16.7)	0.07
Diabetes/ Heart disease/ high blood pressure/ Severe obesity	266(11.6)	2 (11.8)	258 (11.7)	6 (9.1)	0.08
Kidney disease on dialysis/ liver disease/ cancer/ autoimmune disorder/ neurologic disease	161 (7.0)	1 (5.9)	154 (7.0)	6 (9.1)	0.08
Job role					0.45
Registered nurse or nurse manager	1077 (44.2)	8 (38.1)	1035 (44.2)	34 (48.6)	
Clinician (MD, MD equivalent, APP, trainee)	804 (33.0)	8 (38.1)	780 (33.3)	16 (22.9)	
Research/administrative	115 (4.7)	1 (4.8)	113 (4.8)	1 (1.4)	
Support service	112 (4.6)	0 (0.0)	110 (4.7)	2 (2.9)	
Assistant or phlebotomist	84 (3.4)	1 (4.8)	80 (3.4)	3 (4.3)	
Laboratory or pharmacist	77 (3.2)	1 (4.8)	73 (3.1)	3 (4.3)	
Care coordination	71 (2.9)	0 (0.0)	65 (2.8)	6 (8.6)	
Clinic manager or ward clerk	45 (1.8)	2 (9.5)	42 (1.8)	1 (1.4)	
Respiratory or speech therapist	33 (1.4)	0 (0.0)	31 (1.3)	2 (2.9)	
Environmental/food services	17 (0.7)	0 (0.0)	15 (0.6)	2 (2.9)	
Work duties					0.30
Direct patient care involved in intubating or suctioning patient airways	701 (30.2)	6 (35.3)	670 (30.0)	25 (37.9)	
Direct patient care but not performing any airway procedures	1220 (52.6)	10 (58.8)	1178 (52.7)	32 (48.5)	
Staff with indirect patient contact (e.g., reception, environmental	128 (5.5)	0 (0.0)	127 (5.7)	1 (1.5)	

services)					
Laboratory	58 (2.5)	0 (0.0)	55 (2.5)	3 (4.5)	
Work in healthcare but not with patients or biological samples	81 (3.5)	1 (5.9)	78 (3.5)	2 (3.0)	
Other	131 (5.6)	0 (0.0)	128 (5.7)	3 (4.5)	
Performed AGP[†]	385 (16.0)	3 (18.8)	368 (15.9)	14 (20.0)	0.11
COVID-19 exposure at work					0.28
No exposure	942 (38.9)	8 (47.1)	908 (38.9)	26 (37.1)	
Low risk exposure[‡]	680 (28.1)	6 (35.3)	661 (28.3)	13 (18.6)	
High risk exposure[§]	797 (32.9)	3 (17.6)	763 (32.7)	31 (44.3)	
Time spent in the healthcare workplace					0.26
0 hours/week	28 (1.2)	0 (0.0)	28 (1.2)	0 (0.0)	
<10 hours/week	90 (3.7)	2 (11.8)	85 (3.6)	3 (4.3)	
10-20 hours/week	169 (7.0)	1 (5.9)	165 (7.1)	3 (4.3)	
21-30 hours/week	296 (12.2)	0 (0.0)	287 (12.3)	9 (12.9)	
31-40 hours/week	1148 (47.5)	5 (29.4)	1104 (47.3)	39 (55.7)	
>40 hours/week	688 (28.4)	9 (52.9)	663 (28.4)	16 (22.9)	
Time spent providing direct patient-facing care					0.38
0 hours/week	234 (9.7)	1 (5.9)	226 (9.7)	7 (10.0)	
<10 hours/week	305 (12.6)	3 (17.6)	299 (12.8)	3 (4.3)	
10-20 hours/week	400 (16.5)	4 (23.5)	386 (16.6)	10 (14.3)	
21-30 hours/week	397 (16.4)	0 (0.0)	381 (16.3)	16 (22.9)	
31-40 hours/week	825 (34.1)	5 (29.4)	791 (33.9)	29 (41.4)	
>40 hours/week	258 (10.7)	4 (23.5)	249 (10.7)	5 (7.1)	
Number in household (mean (SD))	2.3 (17.2)	1.5 (1.1)	2.3 (1.7)	2.3 (1.9)	0.02
Any children under the age of 18 in household	841 (36.5)	2 (11.8)	817 (36.8)	22 (33.8)	0.06
Any adults 65 years or older in household	256 (11.1)	1 (5.9)	245 (11.0)	10 (15.4)	0.13
Extent of avoiding contact with people who live outside of your home					0.34
All of the time	80 (3.3)	2 (11.8)	72 (3.1)	6 (8.6)	
Most of the time. I only leave my home to buy food or other essentials or to walk/exercise	1232 (50.9)	8 (47.1)	1197 (51.3)	27 (38.6)	
Some of the time. I have reduced the amount of time I am in public spaces, social gatherings, or at work	1096 (45.3)	7 (41.2)	1052 (45.1)	37 (52.9)	
None of the time	11 (0.5)	0 (0.0)	11 (0.5)	0 (0.0)	
Frequency of mask wearing when leaving home					0.25
All of the time	1778 (73.5)	10 (58.8)	1719 (73.7)	49 (70.0)	
Most of the time	621 (25.7)	7 (41.2)	596 (25.6)	18 (25.7)	
Some of the time	17 (0.7)	0 (0.0)	14 (0.6)	3 (4.3)	
Never	3 (0.1)	0 (0.0)	3 (0.1)	0 (0.0)	
Had community exposure with a person who tested positive for COVID-19	176 (7.3)	1 (5.9)	158 (6.8)	17 (24.3)	0.50
Spent time in another country/state	825 (34.1)	3 (17.6)	798 (34.2)	24 (34.3)	<0.01

Numbers are in N (%) unless otherwise stated.

* Standardized mean difference (SMD) is a comparative measure of effect size between groups. Magnitude of effect is considered to be small if SMD = 0.2, medium if SMD = 0.5, and large if SMD = 0.8.

† Aerosol-generating procedures (AGPs): Intubation, extubation, chest physiotherapy, non-invasive ventilation, open suction, nebulized medications, manual ventilation, bronchoscopy, pulmonary function tests, high frequency ventilation, laryngoscopy, autopsy, cardiopulmonary resuscitation, tracheostomy, or high flow nasal cannula use.

‡ Low-risk exposure at work: interacting with a COVID-19 patient with no reported breach in PPE or other safety protocols.

§ High-risk exposure at work: ever interacting with a COVID-19 patient without full PPE or with having a breach in any safety protocol.

Abbreviations: APP = Advanced practice providers, AGP = Aerosol generating procedures, COPD =Chronic Obstructive Pulmonary Disease, MD = Medical Doctor, SD = Standard Deviation, SMD = Standardized mean difference, SHC = Stanford Health Care, UCSF = University of California, San Francisco, ZSFGH = Zuckerberg San Francisco General Hospital

Table 2. Marginal Structural Model of Variables Associated with Incident COVID-19

Self-reported Time-varying Variable	Adjusted Hazard Ratio (95% CI)	p-value
COVID-19 exposure at work		0.013
No exposure	Ref.	
Low risk[A]	0.8 (0.4, 1.7)	
High risk[B]	2.5 (1.3, 4.8)	
Time spent in the healthcare workplace		0.68
<10 hours/week	Ref.	
10-20 hours/week	1.9 (0.5, 7.9)	
21-30 hours/week	2.5 (0.7, 8.9)	
31-40 hours/week	2.2 (0.7, 7.0)	
>40 hours/week	2.0 (0.6, 6.6)	
Time spent providing direct patient-facing care		0.21
<10 hours/week	0.5 (0.2, 0.9)	
10-20 hours/week	0.6 (0.3, 1.3)	
21-30 hours/week	0.9 (0.5, 1.7)	
31-40 hours/week	Ref.	
>40 hours/week	0.9 (0.4, 2.0)	
Direct contact with a person who tested positive for COVID-19 outside of place of work (yes vs. no)	8.1 (3.8, 17.5)	<0.001
Extent of avoidance of people who live outside of your home when not at work (all/most of the time vs. some/none)	1.0 (0.6, 1.6)	0.91
Mask adherence when not at work (all of the time vs. most/some/never)	0.8 (0.5, 1.6)	0.59
Average number of new COVID-19 cases per day in 6 Bay Area counties over 14 days prior to the visit (1 unit increase per 10,000 cases)	1.3 (0.97, 1.8)	0.08

For each marginal structural model, we estimated inverse probability of treatment weights in which exposure probability was estimated for each participant conditioning on fixed variables (age, gender, race, ethnicity, household size, role, and work category) and time-varying variables shown in the table.

[A] Low-risk exposure at work was defined as: ever providing direct care, being within 6 feet, or being within the environment of a COVID-19 patient or being involved with laboratory processing of samples from COVID-19 patients.

[B] High-risk exposure at work was defined as: ever interacting with a COVID-19 patient without full PPE or having a breach in PPE (e.g., tears, accidental removal).

Figure 1. CONSORT diagram

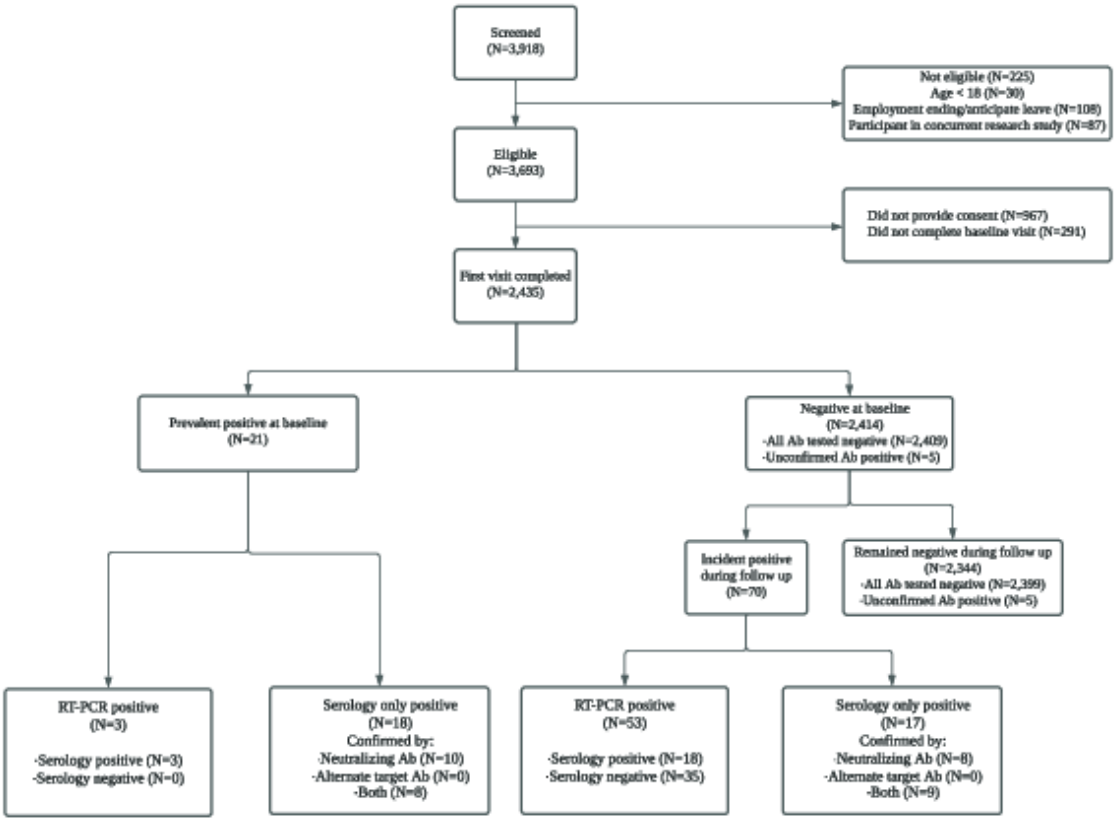
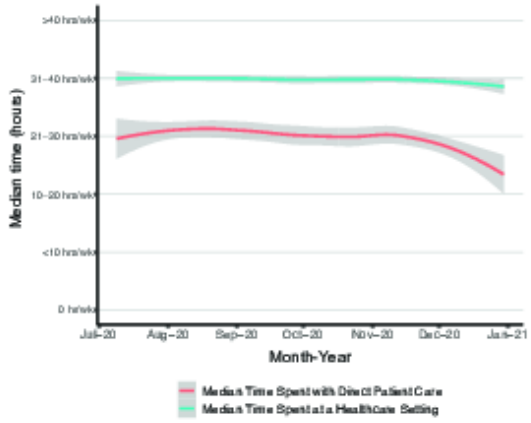
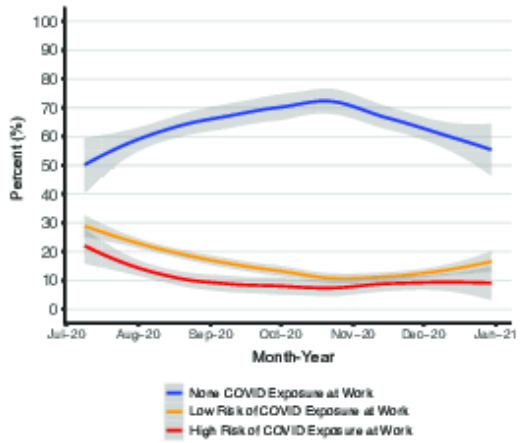


Figure 2. Healthcare worker COVID-19 exposures and incident cases over time

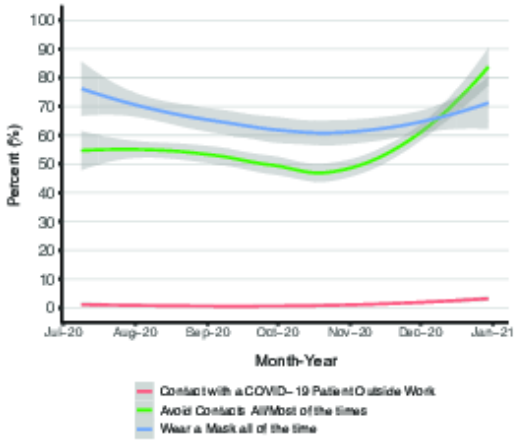
A. Time spent in healthcare setting or in direct patient care



B. Composite COVID-19 exposures at work



C. COVID-19 exposures outside of work



D. Community and healthcare worker COVID-19 infections over time

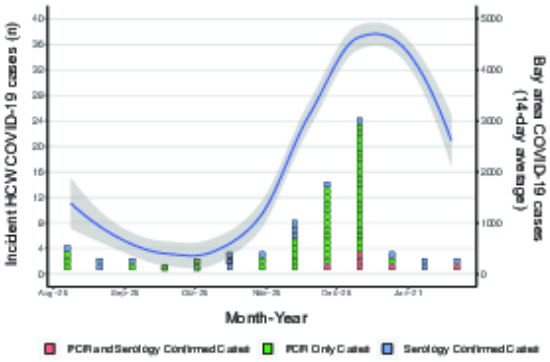


Figure 3. Healthcare worker COVID-19 exposures and incident cases over time