medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Risk of symptomatic Covid-19 among frontline healthcare workers

Long H. Nguyen, M.D., M.S.[1,2,3]*, David A. Drew, Ph.D.[1,2]*, Amit D. Joshi, MBBS, Ph.D.[1,2], Chuan-Guo Guo[1,2,4], M.S., Wenjie Ma, MBBS, Sc.D.[1,2,3], Raaj S. Mehta, M.D.[1,2], Daniel R. Sikavi, M.D.[5], Chun-Han Lo, M.D., M.P.H.[1,2,6], Sohee Kwon, M.D., M.P.H.[1,2,3], Mingyang Song, M.D., Sc.D.[1,2,6,7], Lorelei A. Mucci, Sc.D.[6], Meir J. Stampfer, M.D., Dr.P.H.[6,8], Walter C. Willett, M.D., Dr.P.H.[6,7], A. Heather Eliassen, Sc.D.[6], Jaime E. Hart, Sc.D.[8,9], Jorge E. Chavarro, M.D., Sc.D.[6,7,8], Janet W. Rich-Edwards, Sc.D.[6,10], Christina C. Lawson, Ph.D.[11], Richard Davies, M.A.[12], Joan Capdevila Pujol, Ph.D.[12], Karla A. Lee, MBBCh, M.Sc.[13], Mary Ni Lochlainn, MBBCh[13], Thomas Varsavsky, M.Sc.[14], Mark S. Graham, Ph.D.[14], Carole H. Sudre, Ph.D.[14], M. Jorge Cardoso, Ph.D.[14], Jonathan Wolf, B.Sc.[12], Sebastien Ourselin, Ph.D.[14], Claire J. Steves, Ph.D.[13], Tim D. Spector, M.D.[13], Andrew T. Chan, M.D., M.P.H.[1,2,15,16] on behalf of the COPE Consortium

*Contributed equally to this work

1. Division of Gastroenterology, Massachusetts General Hospital and Harvard Medical School. Boston, MA, USA.
2. Clinical and Translational Epidemiology Unit, Massachusetts General Hospital and Harvard Medical School. Boston, MA, USA.
3. Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
4. Department of Medicine, Li Ka Shing Faculty of Medicine, University of Hong Kong, Hong Kong, China.
5. Department of Medicine, Massachusetts General Hospital and Harvard Medical School, Boston, MA, USA.
6. Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
7. Department of Nutrition, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
8. Channing Division of Network Medicine, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA.
9. Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, MA, USA.
10. Division of Women's Health, Department of Medicine, Brigham and Women's Hospital and Harvard Medical School. Boston, MA, U.S.A.
11. National Institute for Occupational Safety and Health. Cincinnati, OH. USA.
12. Zoe Global Limited. London, U.K.
13. Department of Twin Research and Genetic Epidemiology, King's College London, London, U.K.
14. School of Biomedical Engineering & Imaging Sciences, King's College London. London, U.K.
15. Department of Immunology and Infectious Disease, Harvard T.H. Chan School of Public Health. Boston, MA, USA.
16. Broad Institute of MIT and Harvard. Cambridge, MA, USA.

Correspondence:
Andrew T. Chan, M.D., M.P.H.
Massachusetts General Hospital and Harvard Medical School
100 Cambridge Street, 15th Floor
Boston, MA 02114
achan@mgh.harvard.edu

Funding/Grant Support:
Zoe provided in kind support for all aspects of building, running and supporting the tracking app and service to all users worldwide. King's College of London investigators (KAL, MNL, TV, MG, CHS, MJC, SO, CJS, TDS) were supported by the Wellcome Trust and EPSRC (WT212904/Z/18/Z, WT203148/Z/16/Z, WT213038/Z/18/Z), the NIHR GSTT/KCL Biomedical Research Centre, MRC/BHF (MR/M016560/1), the NIHR, and the Alzheimer's Society (AS-JF-17-011). ATC is the Stuart and Suzanne Steele MGH Research Scholar and Stand Up to Cancer scientist. The National Institutes of Health grants related to this project include: UM1 CA186107 (MJS), U01 CA176726 (AHE, WCW), U01 CA167552 (WCW, LAM), U01 HL145386 (JEC), R24 ES028521 (JEC), P30ES000002 (JEH), and a National Institute for Occupational Safety and Health grant contract 200-2017-M-94186 (JEC). An Evergrande COVID-19 Response Fund Award through the Massachusetts Consortium on Pathogen Readiness (MassCPR) supported MGH investigators (LHN, DAD, ADJ, CGG, WM, RSM, DRS, CHL, SK, MS, ATC).

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Background:

Data are limited about severe acute respiratory syndrome coronavirus 2 (SARS CoV-2) infection risk among frontline healthcare workers (HCWs) and whether personal protective equipment (PPE) mitigates this risk.

Methods:

We performed a prospective cohort study among 2,135,190 individuals using the COVID Symptom Tracker smartphone application to collect information among frontline HCWs and the general community since March 24, 2020 (United Kingdom) and March 29, 2020 (United States). We used Cox proportional hazards modeling to estimate multivariate-adjusted hazard ratios (aHRs) of a positive Covid-19 test through April 23, 2020.

Results:

Among 2,035,395 community individuals and 99,795 frontline HCWs, we documented 5,545 incident reports of a positive Covid-19 test over 34,435,273 person-days. Compared to the general community, frontline HCWs had an aHR of 11.6 (95% CI: 10.9 to 12.3) for a positive test. Frontline HCWs reporting inadequate PPE availability had a 23% increased risk of testing positive (aHR 1.23, 95% CI: 1.03 to 1.46). Compared to HCWs reporting adequate PPE who did not care for Covid-19 patients, HCWs caring for patients with documented Covid-19 had aHRs for a positive test of 4.91 (95% CI: 4.11 to 5.86) if they had adequate PPE and 5.94 (95% CI: 4.57 to 7.72) if they had inadequate PPE.

Conclusions:

Frontline HCWs had a significantly increased risk of symptomatic Covid-19 infection, which was highest among HCWs with inadequate access to PPE who cared for Covid-19 patients. However, adequate supplies of PPE did not completely mitigate high-risk exposures (ClinicalTrials.gov NCT04331509).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Introduction

Since its emergence, the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has become a global health threat.[1] As of April 2020, over 2.8 million cases of Covid-19 have been documented worldwide with nearly 200,000 deaths.[2,3] With ongoing community transmission from asymptomatic individuals across numerous international foci,[4-8] the burden of this disease is expected to continue rising over the coming weeks.[8-12] Consequently, there will be an ongoing need for frontline healthcare workers (HCW) in patient-facing roles. Because this work requires close personal exposure to patients with the virus, frontline HCWs are at high risk of infection, which may contribute to further spread.[13,14] Initial regional estimates suggest frontline HCWs may account for 10-20% of all diagnoses,[15-17] which may be an underestimate when compared to other developed and similarly affected nations such as Italy.[18]

Based on experience with other viruses spread by respiratory droplets, the consistent use of disposable personal protective equipment (PPE) is critical to reducing nosocomial transmission.[19,20] Recent guidelines from the Center for Disease Control (CDC) advocate for the use of medical-grade masks and respirators at all times to deliver clinical care.[21] However, global shortages of masks, face shields, and gowns—exacerbated by surging demand and supply chain disruptions—have been documented, leading to efforts to conserve PPE through extended use or reuse and the development of sterilization protocols.[22-24]

Although addressing the needs of frontline HCWs to effectively respond to the Covid-19 pandemic is a high priority,[17] there is a lack of prospective data to inform such efforts. Thus, we conducted a prospective, population-based study using a novel mobile-based application to examine the risk of testing positive for Covid-19 and/or developing symptoms associated with infection among 2,135,190 individuals in the United Kingdom (U.K.) and the United States (U.S.) between March 24 and April 23, 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Methods

Development and deployment of a smartphone application

The COVID Symptom Tracker is a freely available smartphone application developed by Zoe Global Ltd. in collaboration with the Massachusetts General Hospital and King's College London that offers participants a guided interface to report a range of baseline demographic information and comorbidities, daily information on potential symptoms, and Covid-19 testing. Participants are encouraged to use the application daily, even when asymptomatic, to allow for the longitudinal, prospective collection of symptoms and Covid-19 testing results.

Participant recruitment

The application was launched in the U.K. on March 24, 2020 and available in the U.S. beginning on March 29, 2020.[25] Participants were recruited through social media outreach, as well as invitations from the investigators of long-running prospective cohort studies to study volunteers (**Table S1**). At enrollment, participants provided informed consent to the use of aggregated information for research purposes and agreed to applicable privacy policies and terms of use. This study was approved by the Partners Human Research Committee (Protocol 2020P000909) and King's College London Ethics Committee (REMAS ID 18210, LRS-19/20-18210). This protocol is registered with ClinicalTrials.gov (NCT04331509).

Assessment of risk factors, symptoms, and testing

The information collected through the app has been provided in detail.[25] Briefly, upon first use, participants were asked to provide baseline demographic factors and answered separate questions about a series of suspected risk factors for Covid-19 (**Table S2**). On first use and upon daily reminders, participants were asked if they felt physically normal, and if not, their symptoms, including the presence of fever, fatigue, and loss of smell/taste (**Table S3**). Participants were also asked if they had been tested for Covid-19 (yes/no), and if yes, the results (none, negative, pending, or positive).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

At enrollment, individuals were asked if they worked in health care and if yes, whether they had direct patient contact. For our primary analysis, we defined frontline HCWs as participants who reported direct patient care. Among these individuals, we queried whether they had cared for suspected or documented Covid-19-infected patients and the frequency with which they used PPE (always, sometimes, never). We asked if they had enough PPE when needed, if they had to reuse PPE, or if they did not have enough because of shortages. We classified availability of PPE as adequate if they never required PPE or if they reported always having the PPE they needed (including reused PPE). We classified PPE availability as inadequate if they reported they did not have enough PPE or if it was not available. We also asked HCWs to report the site of their patient care.

Statistical Analysis

Follow-up time started at the time participants first reported on the app and accrued until the report of a positive Covid-19 test or the time of last data entry, whichever occurred first. We employed Cox proportional hazards modeling stratified by age, date, and country to estimate age- and multivariable-adjusted hazard ratios (aHRs) and their 95% confidence intervals (95% CIs). CI widths were not adjusted for multiplicity. A test of correlation between Schoenfeld residuals and survival time demonstrated no violation of the proportional hazards assumption. Covariates were selected *a priori* based on putative risk factors and included sex, history of diabetes, heart disease, lung disease, kidney disease, current smoking status (each yes/no), and body mass index (17-19.9, 20-24.9, 25-29.9, and ≥30 kg/m^2). Missing categorical data was denoted with a missing indicator, while imputation methods for missing numerical data replaced no more than 5% with the median.

Because the primary outcome (report of a positive Covid-19 test) was dependent on receiving a test, we performed several secondary analyses to ensure the robustness of our findings. First, we leveraged a symptom-based classifier developed by our group that is predictive of positive Covid testing.[26] Briefly, using logistic regression and symptoms preceding confirmatory testing, we found that loss of smell/taste, fatigue, persistent cough, and loss of appetite may predict Covid-19 positivity with high specificity (**Suppl. Methods**). Second, to account for country-specific predictors of obtaining testing, we performed separate inverse probability weighting (IPW) in the U.S. and the U.K. as a function of frontline HCW status and other factors,

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

such as age and symptom burden, followed by inverse probability weighted Cox proportional hazards modeling stratified by 5-year age group and date with additional adjustment for the covariates used in prior models (**Suppl. Methods**). In analyses limited to frontline HCWs, we examined PPE availability and contact with suspected or documented Covid-19 patients, as well as the primary site of clinical practice. Two-sided p-values <0.05 were considered statistically significant. All analyses were performed using R 3.6.1 (Vienna, Austria).

Results

Study population

Between March 24 and April 23, 2020, we enrolled 2,810,103 users (2,790,150 in the U.K. and 195,140 in the U.S.), defined as consenting participants who provided baseline information about either feeling normal or having symptoms (**Fig. S1**). Among users, 134,885 (4.8%) reported being a frontline HCW. We found a reported prevalence of 2,747 Covid-19-positive cases per 100,000 frontline HCWs compared to 242 per 100,000 in the general community (**Fig. 1A**). Higher infection rates were reported in New York, New Jersey, and Louisiana in the U.S. and in the areas around London and the Midlands in the U.K. (**Fig. 1B**).

After excluding 670,298 participants who had follow-up time of less than 24 hours and 4,615 who reported a baseline positive Covid-19 test, we included 2,135,190 participants in our prospective inception cohort, among whom 99,795 persons identified as frontline HCWs (**Fig. S1**). In this cohort, we recorded 24.4 million entries or 11.4 logs per participant with a median follow-up of 18.9 days (interquartile range [IQR] 5.1 to 26.1). The median age was 44 years (IQR 32 to 57). Compared to the general community, frontline HCWs were more frequently female, had a slightly higher prevalence of BMI ≥30, were slightly more likely to smoke, and use several common medications (**Table 1** & **Table S3**). At baseline, 20.2% of frontline HCWs reported at least one symptom associated with Covid-19 infection compared to 14.4% of the general population with fatigue, loss of smell/taste, and hoarse voice being particularly more frequent (**Table S4**).

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Risk of positive Covid-19 testing and symptoms in HCWs

We documented 5,545 incident reports of positive Covid-19 testing over 34,435,273 person-days. Compared with the general community, frontline HCWs had a 12-fold increase in risk of a positive test after multivariable adjustment (aHR 11.6, 95% CI: 10.9 to 12.3; **Table 2** and **Fig. S2**). The association appeared stronger in the U.K. (aHR 12.5, 95% CI: 11.7 to 13.2) compared to the U.S. (aHR 2.80, 95% CI: 2.09 to 3.75). An analysis according to sex demonstrated similar risk estimates among male (aHR 14.0, 95% CI: 12.4 to 15.8) and female (aHR 11.3, 95% CI: 10.5-12.1) frontline HCWs.

We considered the possibility that the observed difference in risk in the U.K. vs. the U.S. may be related to differences in the risk profile of individuals eligible for testing. A multivariable-adjusted IPW Cox proportional hazards model weighted by predictors of testing demonstrated consistent associations between being a frontline HCW and reporting a positive Covid-19 test in the U.S. (aHR 1.29, 95% CI: 1.19 to 1.41) and the U.K. (aHR 1.92, 95% CI: 1.89 to 1.94). In a secondary analysis, we used a validated model based on a combination of symptoms to predict likely Covid-19 infection.[26] Compared with the general community, HCWs initially free of symptoms had an aHR of 1.74 (95% CI: 1.71 to 1.77) for predicted Covid-19, which was similar in the U.S. (aHR 1.26, 95% CI: 1.18 to 1.34) and U.K (aHR 1.79, 95% CI: 1.76 to 1.82).

PPE usage in frontline HCWs

Among frontline HCWs, we assessed the availability of PPE in relation to exposure to Covid-19 patients and subsequent risk for testing positive. Overall, 86% of frontline HCWs reported adequate PPE availability (new and reused), including 85% of those caring for documented Covid-19-positive patients. Compared to those reporting adequate availability, frontline HCWs reporting inadequate PPE had a 23% increased risk of subsequently testing positive for Covid-19 (aHR 1.23, 95% CI: 1.03 to 1.46; **Table 3**). Frontline HCWs with inadequate PPE in direct contact with a documented Covid-19 positive patient had an aHR of 5.94 (95% CI: 4.57 to 7.72) for a positive Covid test compared to those with adequate PPE who were not in contact with suspected or documented Covid-19 patients. Estimates were not materially altered when availability of new

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

and reused PPE were analyzed separately (data not shown). Notably, even among those reporting adequate PPE availability, the aHR for a positive Covid-19 test was 2.54 (95% CI: 2.06 to 3.13) for those caring for suspected Covid-19 patients and 4.91 (95% CI: 4.11 to 5.86) for those caring for documented Covid-19 patients compared to HCWs who did not care for either group.

Workplace location

We examined whether the elevated risk of a positive Covid-19 test differed according to practice location. Compared to the general community, the aHRs for frontline HCWs were 24.3 (95% CI: 21.8 to 27.1) for those working in inpatient settings; 16.2 (95% CI: 13.4 to 19.7) for nursing homes; 11.2 (95% CI: 8.44 to 14.9) for hospital-based clinics; 7.86 (95% CI: 5.63 to 11.0) for home health sites; 6.94 (95% CI: 5.12 to 9.41) for free-standing ambulatory clinics; and 9.52 (95% CI: 7.49 to 12.1) for all others (**Table 4**).

Discussion

Among 2,135,190 participants in the U.K. and U.S. assessed between March 24, 2020 and April 23, 2020, we found that frontline HCWs had up to a 12-fold increased risk of a positive Covid-19 test compared to members of the general community, even after accounting for other risk factors for infection. Among frontline HCWs, inadequate PPE availability was associated with a subsequent 23% increased risk of Covid-19. Although HCWs caring for Covid-19 patients who reported inadequate PPE availability had the highest risk, an increased susceptibility to infection was evident even with adequate PPE. Frontline HCWs who worked in inpatient settings and nursing homes had the greatest risk.

Our findings complement early data showing that nearly 20% of the currently documented cases of Covid-19 are among HCWs. Our results provide a more precise assessment of the magnitude of increased risk among HCWs compared to their surrounding communities in the initial phases of this pandemic in both the U.K. and U.S. We also offer supportive evidence that sufficient availability of PPE, including reused PPE, reduces the risk of Covid-19 infection. However, even with reportedly adequate PPE availability, HCWs who cared for patients with documented Covid-19 remained at elevated risk, highlighting the importance of ensuring not only

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

PPE availability, but also appropriate usage. In addition, the apparent lack of complete protection against acquiring Covid-19 infection through adequate supplies of PPE suggests that additional risk mitigation strategies require further investigation.

Our results are supported by historical data during similar infectious disease outbreaks. During the Ebola crisis, a disease with a comparable reproduction number (e.g. the R_0 measure of how many new cases can be generated from one infected individual), HCWs comprised 3.9% of all cases, an incidence 21-to-32-times greater than the general public.[27,28] During the severe acute respiratory syndrome coronavirus (SARS-CoV or SARS) epidemic, HCWs comprised 20-40% of all cases,[29-32] and inadequate PPE availability was associated with an increased risk of illness among HCWs.[32] The experience with influenza A virus subtype H1N1 reaffirmed the importance of PPE[33,34] and also showed much higher infection rates among HCWs working in dedicated infection containment units.[35]

The strengths of this study include the use of a mobile application to rapidly collect prospective data from a large multinational cohort, which offers immediately actionable risk estimates to inform the public health response to an ongoing pandemic.[36] By recruiting participants through existing cohort studies (https://www.monganinstitute.org/cope-consortium), these results also provide proof-of-concept of the feasibility of leveraging existing infrastructure and engaged participants to address a key knowledge gap. Second, we collected information from participants initially free of a positive Covid-19 test, which offered an opportunity to prospectively assess risk factors for incident symptomatic infection with minimal recall bias. Third, our study design documented initial onset of symptoms, which minimizes biases related to capturing only more severe cases through hospitalization records or death reports. Finally, we collected information on a wide range of known or suspected risk factors for Covid-19 infection which are generally not available in existing registries or population-scale surveillance efforts.

We acknowledge several limitations. First, full details of some exposures were limited to ensure our survey instrument was brief, particularly for time-limited HCWs. For example, we did not ask about specific

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

occupations, experience level, or receipt of PPE training (e.g. mask fit-testing or donning and doffing). Second, our findings are based on self-report. Given the quickened pace of the pandemic, self-reported information was more feasible to collect. In future studies, self-reports with linkage to other sources (e.g. electronic health records) may be possible. Third, our cohort is not a random sampling of the population. Although this limitation is inherent to any study requiring voluntary provision of health information, we acknowledge that data collection through smartphone adoption has comparatively lower penetrance among certain socioeconomic groups, as well as older adults, despite being used by 81% of the U.S. adult population.[37] In future studies, we plan more targeted outreach of populations under-represented in the current cohort, as well as additional data collection instruments (e.g. web or phone surveys) that may be more accessible to certain groups. Finally, our primary outcome is based on report of a positive Covid-19 test. During the time frame of this study, this would generally reflect nasal or throat swabs, which are reasonably specific, but only modestly sensitive.[38] However, any testing misclassification should be non-differential by occupation. Our results may also reflect differences in access to testing among HCWs compared to the general community, as well as true differences in infection rates. In our secondary analyses, we also found that HCWs were at higher risk of a positive test even after adjusting for greater probability of receiving a test. Furthermore, HCWs were at higher risk of developing symptoms that predicted Covid-19, which does not reflect access to testing. Ideally, we would assess the risk of Covid-19 infection within a population which has undergone uniform screening. However, the current shortage of PCR-based testing kits does not make such an approach feasible. Future studies using serologic testing to ascertain Covid-19 infection will require additional assessment of test performance and the ability to distinguish recent or active infection from past exposure.

In conclusion, within a large population-based sample of 2,135,190 individuals in the U.S. and U.K., we demonstrate a significantly increased risk of Covid-19 infection among frontline HCWs compared to the general community. This risk is greatest among individuals in direct contact with Covid-19 patients who report inadequate PPE availability, supporting the importance of providing adequate supplies. However, because infection risk remained elevated even with adequate PPE, our results suggest the need to ensure the proper use of PPE as well as adherence to other infection control measures.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Acknowledgements

We would like to thank the more than 2.9 million contributing citizen researchers who have downloaded the COVID Symptom Tracker, including participants of cohort studies within the Coronovirus Pandemic Epidemiology (COPE) Consortium (**Table S4**). We thank the investigators of the cohort studies enrolled in the COPE Consortium for their assistance in disseminating the COVID Symptom Tracker to their study participants. We also thank the Stand Up to Cancer Foundation for their assistance in media and social media outreach. The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the National Institute for Occupational Safety and Health, Centers for Disease Control and Prevention.

Conflicts of interest

JW, RD, and JC are employees of Zoe Global Ltd. TDS is a consultant to Zoe Global Ltd. DAD and ATC previously served as investigators on a clinical trial of diet and lifestyle using a separate mobile application that was supported by Zoe Global Ltd. Other authors have no conflict of interest to declare.

Data availability

Data collected in the app is being shared with other health researchers through the NHS-funded Health Data Research U.K. (HDRUK)/SAIL consortium, housed in the U.K. Secure Research Platform (UKSeRP) in Swansea. Anonymized data is available to be shared with bonafide researchers HDRUK according to their protocols in the public interest (https://healthdatagateway.org/detail/9b604483-9cdc-41b2-b82c-14ee3dd705f6). U.S. investigators are encouraged to coordinate data requests through the COPE Consortium (www.monganinstitute.org/cope-consortium). Data updates can be found on https://covid.joinzoe.com.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

REFERENCES

1. Anderson RM, Heesterbeek H, Klinkenberg D, Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? Lancet (London, England) 2020;395:931-4.
2. WHO Coronavirus disease (COVID-2019) Situation Reports. at https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/.)
3. Johns Hopkins University of Medicine Coronavirus Resource Center. 2020. at https://coronavirus.jhu.edu/us-map.)
4. Wilder-Smith A, Chiew CJ, Lee VJ. Can we contain the COVID-19 outbreak with the same measures as for SARS? Lancet Infect Dis 2020.
5. Zou L, Ruan F, Huang M, et al. SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. N Engl J Med 2020;382:1177-9.
6. Bai Y, Yao L, Wei T, et al. Presumed Asymptomatic Carrier Transmission of COVID-19. Jama 2020.
7. Arons MM, Hatfield KM, Reddy SC, et al. Presymptomatic SARS-CoV-2 Infections and Transmission in a Skilled Nursing Facility. N Engl J Med 2020.
8. Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet (London, England) 2020;395:1054-62.
9. Remuzzi A, Remuzzi G. COVID-19 and Italy: what next? Lancet (London, England) 2020;395:1225-8.
10. Wang D, Hu B, Hu C, et al. Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. Jama 2020.
11. Grasselli G, Pesenti A, Cecconi M. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. Jama 2020.
12. Guan WJ, Ni ZY, Hu Y, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med 2020.
13. Lau H, Khosrawipour V, Kocbach P, et al. The positive impact of lockdown in Wuhan on containing the COVID-19 outbreak in China. J Travel Med 2020.
14. Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. Jama 2020.
15. California Department of Public Health Latest Covid-19 Facts. 2020. (Accessed April 20, 2020, at https://www.cdph.ca.gov/Programs/OPA/Pages/NR20-065.aspx.)
16. Wisconsin Department of Health Services Covid-19: Wisconsin Cases. 2020. (Accessed April 21. 2020, at https://www.dhs.wisconsin.gov/covid-19/cases.htm.)
17. Team CC-R. Characteristics of Health Care Personnel with COVID-19 - United States, February 12-April 9, 2020. MMWR Morb Mortal Wkly Rep 2020;69:477-81.
18. Lazzerini M, Putoto G. COVID-19 in Italy: momentous decisions and many uncertainties. Lancet Glob Health 2020.
19. Verbeek JH, Rajamaki B, Ijaz S, et al. Personal protective equipment for preventing highly infectious diseases due to exposure to contaminated body fluids in healthcare staff. Cochrane Database Syst Rev 2020;4:CD011621.
20. Tran K, Cimon K, Severn M, Pessoa-Silva CL, Conly J. Aerosol generating procedures and risk of transmission of acute respiratory infections to healthcare workers: a systematic review. PLoS One 2012;7:e35797.
21. Center for Disease Control and Prevention: Interim Infection Prevention and Control Recommendations for Patients with Suspected of Confirmed Coronavirus Disease 2019 (COVID-19) in Health Care Settings. (Accessed April 13, 2020, at https://www.cdc.gov/coronavirus/2019-ncov/hcp/infection-control-recommendations.html.)
22. Fischer R, Morris DH, van Doremalen N, et al. Assessment of N95 respirator decontamination and re-use for SARS-CoV-2. medRxiv 2020:2020.04.11.20062018.
23. Schwartz A, Stiegel M, Greeson N, et al. Decontamination and Reuse of N95 Respirators with Hydrogen Peroxide Vapor to Address Worldwide Personal Protective Equipment Shortages During the SARS-CoV-2 (COVID-19) Pandemic. Applied Biosafety 2020:1535676020919932.
24. Livingston E, Desai A, Berkwits M. Sourcing Personal Protective Equipment During the COVID-19 Pandemic. Jama 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

25. Drew DA, Nguyen LH, Steves CJ, Wolf J, Spector TD, Chan AT. Rapid implementation of mobile technology for real-time epidemiology of COVID-19. medRxiv 2020:2020.04.02.20051334.
26. Menni C, Valdes A, Freidin M, et al. Loss of smell and taste in combination with other symptoms is a potential predictor of COVID-19 infection (in press) Nature Medicine 2020.
27. Health worker Ebola infections in Guinea, Liberia, and Sierra Leone. 2015. (Accessed April 15, 2020,
28. Houlihan CF, McGowan CR, Dicks S, et al. Ebola exposure, illness experience, and Ebola antibody prevalence in international responders to the West African Ebola epidemic 2014-2016: A cross-sectional study. PLoS Med 2017;14:e1002300.
29. Koh D, Lim MK, Chia SE. SARS: health care work can be hazardous to health. Occup Med (Lond) 2003;53:241-3.
30. Seto WH, Tsang D, Yung RW, et al. Effectiveness of precautions against droplets and contact in prevention of nosocomial transmission of severe acute respiratory syndrome (SARS). Lancet (London, England) 2003;361:1519-20.
31. Loeb M, McGeer A, Henry B, et al. SARS among critical care nurses, Toronto. Emerg Infect Dis 2004;10:251-5.
32. Lau JT, Fung KS, Wong TW, et al. SARS transmission among hospital workers in Hong Kong. Emerg Infect Dis 2004;10:280-6.
33. Marshall C, Kelso A, McBryde E, et al. Pandemic (H1N1) 2009 risk for frontline health care workers. Emerg Infect Dis 2011;17:1000-6.
34. Cheng VC, Tai JW, Wong LM, et al. Prevention of nosocomial transmission of swine-origin pandemic influenza virus A/H1N1 by infection control bundle. J Hosp Infect 2010;74:271-7.
35. Chen MI, Lee VJ, Barr I, et al. Risk factors for pandemic (H1N1) 2009 virus seroconversion among hospital staff, Singapore. Emerg Infect Dis 2010;16:1554-61.
36. Brownstein JS, Freifeld CC, Madoff LC. Digital disease detection--harnessing the Web for public health surveillance. N Engl J Med 2009;360:2153-5, 7.
37. Pew Research Center for Internet & Technology: Mobile Fact Sheet. 2020. (Accessed April 17, 2020, at https://www.pewresearch.org/internet/fact-sheet/mobile/.)
38. Wang W, Xu Y, Gao R, et al. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. Jama 2020.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table 1. Baseline characteristics of frontline healthcare workers compared to the general public.

	Participants (*n*=2,135,190)	
	Frontline HCWs (*n*=99,795)	**General community (*n*=2,035,395)**
Country (%)		
U.S.	14.6	6.1
U.K.	85.4	93.9
Age (years)	42 [33, 53]	44 [33, 56]
<25	4.5	5.0
25-34	24.7	20.4
35-44	25.4	22.8
45-54	33.9	20.7
55-64	17.7	17.2
≥65	3.9	13.9
Male sex (%)	17.0	37.0
Ethnicity		
Hispanic	6.4	4.9
Non-Hispanic	91.5	92.0
Prefer not to say	2.0	3.2
Race		
Black	1.9	1.4
White	92.2	94.6
Asian	4.2	2.1

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Other	1.7	1.8
BMI (kg/m^2)	25.8 [22.8, 30.2]	25.3 [22.5, 29.1]
17-19.9	5.8	8.3
20-24.9	38.1	39.2
25-29.9	30.1	31.5
≥30	25.9	21.1
Comorbidities (%)		
Diabetes	2.5	3.1
Heart Disease	1.6	2.4
Lung Disease	13.1	12.2
Kidney Disease	0.6	0.7
Cancer	0.5	1.0
Pregnant (% of females)	0.9	1.0
Medication usage (%)		
NSAIDs	8.2	6.1
Immunosuppressants	2.5	3.2
Chemotherapy/Immunotherapy	0.1	0.3
Aspirin	4.0	5.2
ACE inhibitor	5.0	4.9
Current smoking (%)	9.7	7.6
Housebound (%)	2.5	4.2

Abbreviations: ACE (angiotensin converting enzyme), BMI (body mass index), m (meter), kg (kilogram),

Non-steroidal anti-inflammatory drugs (NSAIDs)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Median [IQR] is presented for continuous variables.

† includes healthcare workers without direct patient care/patient contact responsibilities

Frequencies and proportions are calculated based on the total number of participants with available data.

History of cancer, aspirin use, and smoking status have been queried since launch in the U.S. and since 3/29/2020 in the U.K. Race and ethnicity questions were queried as of 4/17/2020

Table 2. Risk of testing positive for Covid-19 or predicted Covid-19 infection among HCWs compared with the general community

	No. with Event/Person-days	Incidence (30-day)	Hazard Ratio (95% CI)	
			Age-adjusted	Multivariate-adjusted
Positive Covid-19 testing				
General community	3,623/32,980,571	0.33%	1.0 (ref)	1.0 (ref)
Frontline healthcare worker	1,922/1,454,701	3.96%	11.7 (11.0 to 12.4)	11.6 (10.9 to 12.3)
Predicted Covid-19 infection†				
General community	56,059/29,864,522	5.63%	1.0 (ref.)	1.0 (ref.)
Frontline healthcare worker	5,022/1,242,857	12.1%	2.05 (1.99 to 2.11)	2.04 (1.98 to 2.10)

Abbreviations: CI (confidence interval)

All models were stratified by 5-year age group, calendar date at study entry, and country.

Multivariate risk factor models were adjusted for sex, history of diabetes, heart disease, lung disease, kidney disease, and current smoking (each yes/no), and body mass index (17-19.9, 20-24.9, 25-29.9, and ≥30 kg/m^2).

† Using a symptom-based model described in Menni et al *(Nature Medicine*, in press)

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

Table 3. Risk of testing positive for Covid-19 according to personal protective equipment (PPE) availability and exposure to Covid-19 patients among frontline healthcare workers.

	Personal protective equipment	
	Adequate availability†	**Inadequate availability**
No. with Event/Person-days	738/413,629	157/60,917
Unadjusted HR (95% CI)	1.0 (ref)	1.24 (1.04 to 1.47)
Multivariate-adjusted HR (95% CI)	1.0 (ref)	1.23 (1.03 to 1.46)
Exposure to patients with Covid-19		
None		
No. with Event/Person-days	205/265,253	48/35,159
Unadjusted HR (95% CI)	1.0 (ref)	1.54 (1.12 to 2.11)
Multivariate-adjusted HR (95% CI)	1.0 (ref)	1.53 (1.11 to 2.09)
Suspected Covid-19 patients		
No. with Event/Person-days	162/74,054	26/14,083
Unadjusted HR (95% CI)	2.56 (2.08 to 3.15)	1.88 (1.25 to 2.84)
Multivariate-adjusted HR (95% CI)	2.54 (2.06 to 3.13)	1.84 (1.22 to 2.78)
Documented Covid-19 patients		
No. with Event/Person-days	371/74,322	83/11,675
Unadjusted HR (95% CI)	4.99 (4.18 to 5.96)	5.98 (4.61 to 7.77)
Multivariate-adjusted HR (95% CI)	4.91 (4.11 to 5.86)	5.94 (4.57 to 7.72)

Abbreviations: CI (confidence interval), HR (hazard ratio)

† includes new and reused PPE

All models were stratified by 5-year age group, calendar date at study entry, and country.

Multivariate risk factor models were adjusted for sex, history of diabetes, heart disease, lung disease,

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .

kidney disease, and current smoking (each yes/no), and body mass index (17-19.9, 20-24.9, 25-29.9, and ≥30 kg/m^2).

Table 4. Frontline healthcare workers and risk of testing positive for Covid-19 by site of care delivery.

	Positive Covid-19 testing			
			Hazard Ratio (95% CI)	
	No. with Event/Person-days	Incidence (30-day)	Age-adjusted	Multivariate-adjusted
General population	3,623/32,980,571	0.33%	1.0 (ref)	1.0 (ref)
Frontline healthcare worker				
Inpatient	564/184,293	9.18%	23.6 (21.2 to 26.2)	24.3 (21.8 to 27.1)
Nursing homes	118/52,901	6.69%	16.5 (13.6 to 20.0)	16.2 (13.4 to 19.7)
Outpatient clinics in hospital	51/45,217	3.38%	10.7 (8.10 to 14.3)	11.2 (8.44 to 14.9)
Home health sites	36/38,642	2.79%	7.79 (5.58 to 10.9)	7.86 (5.63 to 11.0)
Ambulatory clinics	44/66,408	1.99%	6.64 (4.90 to 9.01)	6.94 (5.12 to 9.41)
Other	73/64,310	3.41%	9.42 (7.42 to 12.0)	9.52 (7.49 to 12.1)

Abbreviations: CI (confidence interval)

Model was stratified by 5-year age group, calendar date at study entry, and country and adjusted for sex, history of diabetes, heart disease, lung disease, kidney disease, and current smoking (each yes/no), and body mass index (17-19.9, 20-24.9, 25-29.9, and ≥30 kg/m^2).

Ambulatory clinics include free-standing (non-hospital) primary care/specialty clinics and school-based clinics

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license.

medRxiv preprint doi: https://doi.org/10.1101/2020.04.29.20084111; this version posted May 5, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-ND 4.0 International license .
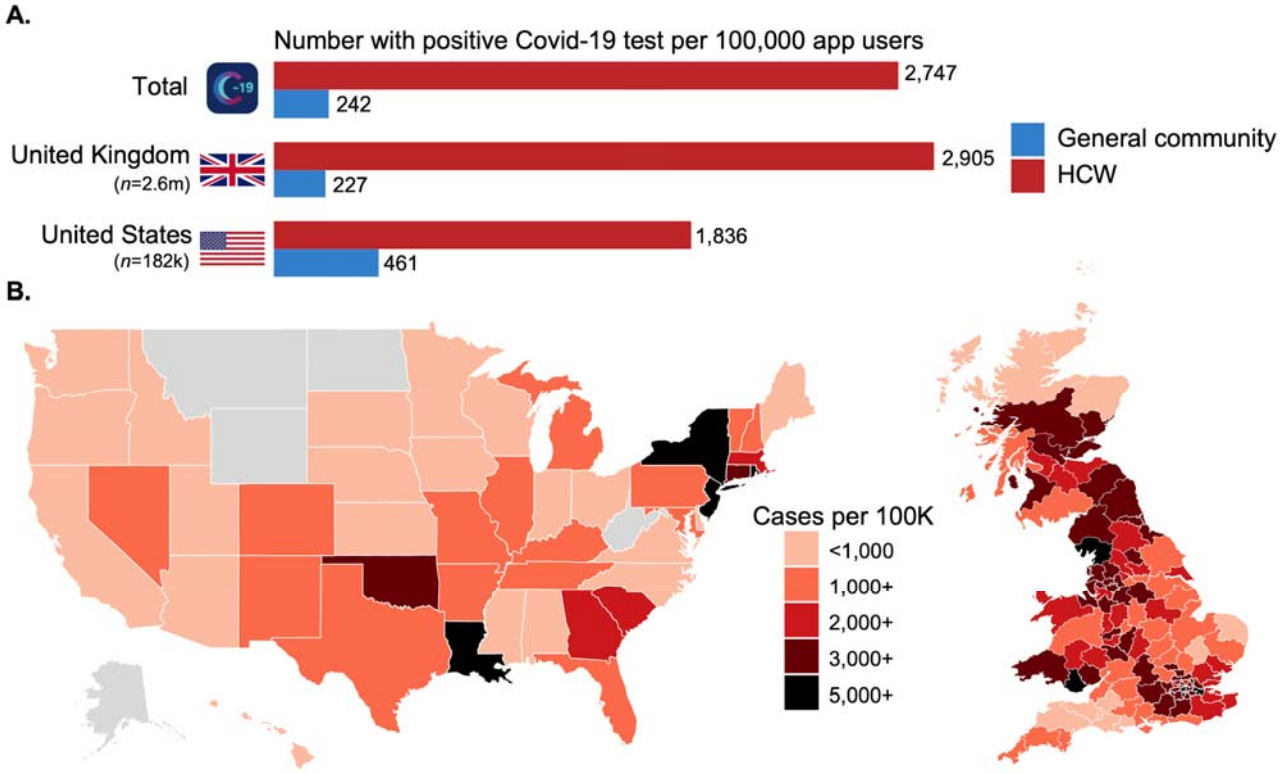
Figure 1. The risk of testing positive for Covid-19 among frontline healthcare workers (HCW). A. Between March 24, 2020 and April 23, 2020, considerable disparities in prevalence of a positive Covid-19 test among frontline HCW risk compared to the general community were observed in both the United Kingdom and the United States. **B**. Prevalence of a positive COVID-19 test reported by frontline HCWs in the United States and the United Kingdom. Regions in gray did not offer sufficient data.