STATE OF WASHINGTON
THURSTON COUNTY SUPERIOR COURT

DEPARTMENT OF LABOR AND INDUSTRIES OF THE STATE OF WASHINGTON, Plaintiff, v. FOWLER NAT D. AND MARY M. DBA FARM BOY DRIVE IN Defendant.	NO. 20-2-02460-34 DECLARATION OF SCOTT W. LINDQUIST, MD. MPH

I, Scott Lindquist, M.D., MPH declare under the penalty of perjury under the laws of the State of Washington that the following is true and correct:

1. I am over the age of eighteen and am otherwise competent to testify. I make these statements on personal knowledge and belief.
2. I am the Deputy Health Officer and State Epidemiologist for Communicable Diseases working within the Office of the State Health Office for the Washington State Department of Health (DOH). I have held this position since 2014.
3. I have a MD degree from the University of Washington and a Master's Degree in Public Health from Harvard University.
4. In my capacity as the State Epidemiologist for Communicable Diseases, I provide leadership in epidemiologic oversight and consultation to DOH programs, other

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

agencies, and local health jurisdictions involved in preventing and reducing communicable diseases. During the novel coronavirus (COVID-19) pandemic my role has been as an infectious diseases subject matter expert in the incident command structure.

5. The World Health Organization has provided extensive information regarding COVID-19, which can be found at https://www.who.int/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19 (last visited December 16, 2020). Excerpts from this site appear in the paragraphs below.

6. According to the WHO, "COVID-19 is the disease caused by a new coronavirus called SARS-CoV-2." Common symptoms of COVID-19 include fever, dry cough, and fatigue. Other symptoms include but are not limited to loss of senses of taste or smell, sore throat, muscle or joint pain, nausea or vomiting, and chills or dizziness. Shortness of breath, confusion, loss of appetite, and high fever are among the severe COVID-19 symptoms.

7. The WHO has described the progression of COVID-19 in infected people. Among those who develop symptoms, most (about 80%) recover from the disease without needing hospital treatment. About 15% become seriously ill and require oxygen and 5% become critically ill and need intensive care.

8. Complications of COVID-19 leading to death may include respiratory failure, acute respiratory distress syndrome (ARDS), sepsis and septic shock, thromboembolism, and/or multiorgan failure, including injury of the heart, liver or kidneys.

9. In rare situations, children can develop a severe inflammatory syndrome a few weeks after infection.

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

10. The WHO notes that "[s]ome people who have had COVID-19, whether they have needed hospitalization or not, continue to experience symptoms, including fatigue, respiratory and neurological symptoms."

11. Certain populations are at greater risk of developing severe COVID-19 symptoms. These include people 60 years of age and older and people with underlying medical conditions including high blood pressure, heart and lung problems, diabetes, cancer, or obesity.

12. COVID-19 presents one of the greatest public health crises to ever face Washington State. Since the first appearance of the virus in the state, DOH has worked closely with the Governor's office, state emergency management agencies, and others to remain informed of factual and scientific developments regarding COVID-19.

13. DOH maintains the Washington Disease Reporting System (WDRS), an electronic disease surveillance system that allows public health staff in the state to enter, process, track, and analyze disease-related data. Local health jurisdictions, such as health care entities, clinics, laboratories, hospitals, and other non-health care sources report COVID-19 cases and outbreaks through the WDRS, which allows secure communication and coordination among state and local health departments.

14. Based on information from the WDRS, DOH publishes a weekly Statewide COVID-19 Outbreak Report. The most recent Outbreak Report, released on December 10, 2020, shows that the most common outbreak scenario reported is in restaurants and food service settings. Exposures and transmissions are most easily accomplished in these settings because people are within six feet of each other for extended periods without a mask on (while eating or drinking). Attached as Exhibit A is a true and correct copy of the December 28, 2020 Statewide COVID-19 Outbreak Report, which is available online at: https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/data-tables/StatewideCOVID-19OutbreakReport.pdf.

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

15. On November 15, 2020, the Governor issued a proclamation prohibiting restaurants and bars from providing indoor dine-in services based on the rise in COVID-19 cases in late October and early November and evidence that transmission of this virus is airborne through "very small droplets called aerosols that are expelled from our mouths when we breathe, talk, sing, vocalize, cough, or sneeze." On December 10, 2020, the Governor issued Proclamation 20-25.9, which extended the restrictions established in Proclamation 20-25.8 to January 4, 2021. On December 30, 2020, the Governor issued Proclamation 20-25.11, which extended the restrictions to January 11, 2021. On January 11, 2021, the Governor issued Proclamation No. 20-25.12, which established the "Healthy Washington: Roadmap to Recovery." Details of the Roadmap to Recovery are set out in the accompanying Declaration of Nicholas Streuli.

16. The advent of SARS-CoV-2 that caused COVID-19 has presented unique challenges. At this time, it is believed to be spread from one person to another when an infected person speaks, coughs, or sneezes or when a person touches a contaminated surface or object and then touches their own mouth, nose, or eyes.

17. According to a study by the Department of Atmospheric Sciences at Texas A&M University, airborne transmission is the dominant route for the spread of COVID-19. (Renyi Zhang, et. al, *identifying airborne transmission as the dominant route for the spread of COVID-19*). The primary route of infection is through respiratory droplets transferred from one person to another in close proximity for a period of time. Being within 6 feet of another person for 15 minutes in a day is considered a "close contact" with a significant potential for transmission of the disease. The disease can also be transmitted by contact with surfaces where respiratory droplets have been deposited and then transferring these materials to the eyes, nose, or mouth. Transmission of the disease may occur as well due to small aerosol droplets that can be suspended in air during heavy respiration, activities such as singing, and certain medical procedures.

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

Aerosol transmission risks are increased when interactions occur indoors, when ventilation is limited, and when more people are brought together. Significantly, it can be transmitted by an infected human, whether or not that person is symptomatic. In other words, a person can actively transmit the virus to other humans and not know that they are potentially spreading a potentially fatal virus. A true and correct copy of the Zhang study is attached hereto as Exhibit B.

18. SARS-CoV-2 is considered a novel virus in that it has not been infecting humans prior to the current outbreak and so there is no community immunity or experience with the disease. The disease also builds up a high viral concentration in the upper respiratory track early in the infection so that people become most infectious just before and during early symptoms. Further, the disease has spread widely in the community so that there is a potential for contact with others who have the infection, perhaps before they know it themselves. Although vaccines have now been approved and others may be approved in the near future, it will be months before the vaccines are widely available or sufficiently distributed to impact transmission of the virus.

19. The primary control of COVID-19 virus is to avoid contact with others and maintain a safe distance when interaction is required. Wearing face coverings and other masks can reduce the potential for respiratory droplets reaching others. Limitations on activities in the community are adopted with the purpose of reducing the number of contacts between people, particularly in situations where individuals may have multiple contacts in enclosed environments. Requirements to wear face coverings protect others from respiratory droplets generated by the wearer and increase safety during necessary times of close proximity to others.

20. In a study published in November 2020, researchers from Stanford and Northwestern Universities studied the extent to which "Points of Interest" (POIs) such as restaurants, grocery stores, and religious establishments contributed to the spread of COVID-19.

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

Chang et al., *Mobility network models of COVID-19 explain inequities and inform reopening*. The study concludes that "a small minority of 'superspreader' points of interest account for a large majority of the infections, and that restricting the maximum occupancy at each point of interest is more effective than uniformly distributing mobility." Chang at al. at 1. More specifically, the Chang study determined that "[c]ertain categories of POIs also contributed far more to infections (for example, full-service restaurants and hotels), although [the authors'] model predicted time-dependent variation in how much each category contributed . . . . For example, restaurants and fitness centers contributed less to the predicted number of infections over time, *probably because of lockdown orders to close these POIs* . . . ." Chang et al. at 3 (emphasis added). In other words, facilities such as full-service restaurants are high risk hazards and powerful contributors to the spread of COVID-19, with reductions in their contributions over time being attributed to lockdowns, i.e., closures. A true and correct copy of the Chang report is attached hereto as Exhibit C.

21. In an additional study published in November 2020, researchers found that COVID-19 "droplet transmission can occur at a distance greater than 6 feet if there is direct air flow from an infected person in an indoor setting." In one restaurant, a subject was infected by someone more than 20 feet away. (Kwon et. al, *Evidence of Long-Distance Droplet Transmission of SARS-CoV-2 by Direct Air Flow in a Restaurant in Korea).* The study further noted that the "attack rate" at the subject restaurant was greater than that of the secondary attack rate among total close contacts and household contacts. A true and correct copy of the Kwon report is attached hereto as Exhibit D.

22. In a study published in September 2020, researchers compared community exposures of subjects with and without the COVID-19 infection. The study concluded that "participants with and without COVID-19 reported generally similar community exposures, with the exception of going to locations with on-site eating and drinking

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

options. Adults with confirmed COVID-19 (case-patients) were approximately twice as likely as were control-participants to have reported dining at a restaurant in the 14 days before becoming ill." Fisher et al., *Community and Close Contact Exposures Associated with COVID-19 Among Symptomatic Adults ≥ 18 Years in 11 Outpatient Health Care Facilities – United States, July 2020*. A true and correct copy of the Fisher Report is attached hereto as Exhibit E.

23. In my opinion, pausing the phased reopening and reinstituting restrictions were necessary measures to address the rise in COVID-19 activity, including hospitalizations and deaths, across the state as of mid-July 2020. Disease transmissibility of a virus can be quantified by its basic reproductive number, or R0 (pronounced R naught). The R0 is the average number of new infections that result from a single infected person in a wholly susceptible population. The effective reproduction number or Re is the number of cases generated in the current state of a population. The R0 and Re can vary not only based on characteristics of a virus but also with the contact rate between people, including physical distancing strategies and other mitigation measures. Unmitigated, the Re of COVID-19 initially experienced in Washington was around 2 to 3.5. This means every infected person likely spread the disease on average to 2 to 3.5 other individuals. Through mitigation efforts, the Re of COVID-19 in Western Washington dipped below 1 in mid-May, meaning the number of people contracting COVID-19 was declining.

24. In September 2020, COVID-19 cases and hospitalizations started increasing in Washington and nationwide. On September 5, 2020, the state's seven-day rolling average for new confirmed cases was 421. By October 5, that number had risen to 578, and on November 5, it had more than doubled, reaching 1400. This trend along with increasing hospitalizations for confirmed COVID-19 demonstrated exponential growth of the outbreak. On November 1, 2020, the DOH modeling team estimated that the Re

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

in western Washington was 1.8 and the Re in eastern Washington was 1.7 which was the highest it had been in either western or eastern Washington since March.

25. As noted, cases of COVID-19 infection have increased dramatically throughout the fall and early winter of 2020. Cases in Thurston County remain at an unsafe level, having increased from 41 new cases during the week ending September 20, 2020, to 452 new cases during the week ending January 10, 2021. This data can be found at the Thurston County COVID-19 Dashboard located at; https://www.thurstoncountywa.gov/phss/Coronavirus/Pages/covid-19-dashboard.aspx. The spread of rapidly increasing COVID cases in our State threatens to overwhelm our state's hospital and medical systems.

I declare under penalty of perjury under the laws of the State of Washington and the United States that the foregoing is true and correct.

DATED this 13th day of January, 2021 at Olympia, Washington.

SCOTT LINDQUIST, MD, MPH

ATTORNEY GENERAL OF WASHINGTON
Labor & Industries Division
7141 Cleanwater Drive SW
PO Box 40121
Olympia, WA 98504-0121
360-586-7707

Exhibit A

Washington State Department of Health

Statewide COVID-19 Outbreak Report

Washington State Department of Health

12/10/2020

To request this document in another format, call 1-800-525-0127. Deaf or hard of hearing customers, please call 711 (Washington Relay) or email civil.rights@doh.wa.gov.

Publication Number 420-286

For inquiries about this report from media, contact the Public Information Desk: doh-pio@doh.wa.gov

Coronavirus disease 2019 (COVID-19) surveillance in Washington State is based upon local health jurisdictions (LHJs) reporting information about COVID-19 cases or outbreaks through the Washington Disease Reporting System (WDRS), which can be updated with new information over time. Given this dynamic nature of COVID-19 surveillance data, outbreaks reported previously may be removed from subsequent reports after investigation. Additionally, there may be a large reporting delay for information pertaining to outbreaks, meaning that the number of reported outbreaks may be underestimated for the most recent weeks.

Summary

This document provides an overview of reported COVID-19 outbreaks in Washington State. This report is based on information in the WDRS database. LHJs are asked to report outbreaks to the Washington State Department of Health (DOH) by creating outbreak events in WDRS for each outbreak. We compute an Outbreak Reference Date to approximate the beginning of each outbreak, using all available information, including symptom onset date and outbreak report date. This date parameter is used in the graphs below.

Healthcare settings, including long-term care facilities (LTCFs), are reported and investigated separately from non-healthcare settings due to unique infection prevention considerations in healthcare settings.

Non-healthcare settings in this report include congregate settings where COVID-19 outbreaks can occur. Examples of non-healthcare congregate settings include agricultural settings, public events, schools, childcare, restaurants, food processing facilities, and prisons.

For the purposes of this report, healthcare settings are defined as: long-term care facilities, hospitals, outpatient settings (dental clinics, etc.), behavioral health facilities, supported living facilities, home healthcare, dialysis centers, or independent senior living facilities. Long-term care facilities include nursing facilities, assisted living facilities, and adult family homes.

To request this document in another format, call 1-800-525-0127. Deaf or hard of hearing customers, please call 711 (Washington Relay) or email civil.rights@doh.wa.gov

COVID-19 Outbreaks in Non-Healthcare Congregate Settings

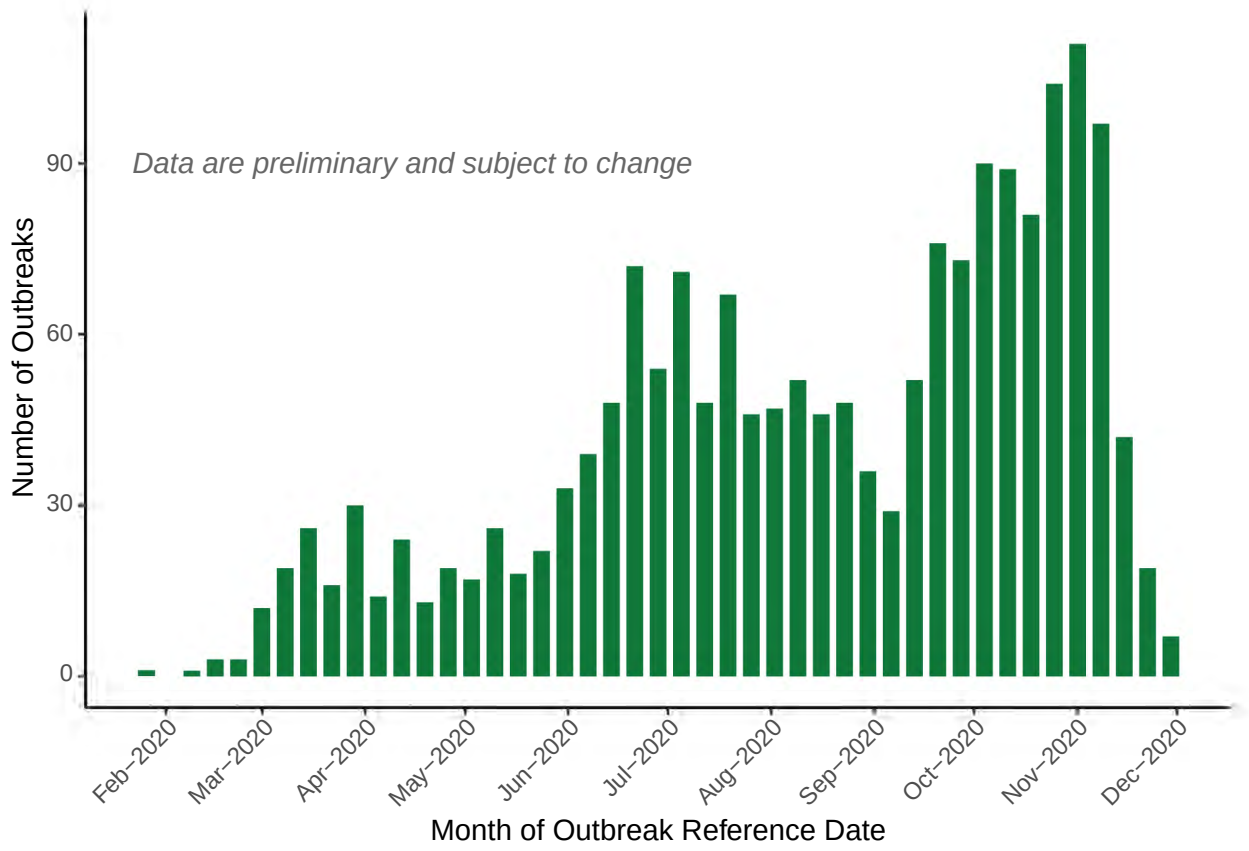
An outbreak in a non-healthcare congregate setting is defined by following criteria being met:

1. Two or more laboratory positive (PCR or antigen) COVID-19 cases, AND
2. At least two of the cases have symptom onset dates within 14 days of each other, AND
3. Plausible epidemiological evidence of transmission in a shared location other than a household is observed

A total of 1841 non-healthcare associated COVID-19 outbreaks have been reported as of Saturday, 12/05/2020. The table below describes type of outbreaks reported by setting. Outbreak reporting prior to June 1st is less complete than more recent weeks. The non-healthcare workplace setting guideline can be found here: https://www.doh.wa.gov/Portals/1/Documents/5100/420-284-Guidance-NonHC-Workplace.pdf

Please note that a backlog of outbreaks in the homeless services category was processed in the last week. This backlog is not reflective of a surge in outbreaks in the homeless service settings in the last week.

Number of non-healthcare congregate settings outbreaks reported by approximate week of initial symptom onset, through 12/05/2020



Number of non-healthcare congregate settings COVID-19 outbreaks reported by setting, ever reported and reported during the most recent week

The most recent week included in this report is Sunday, 11/29/2020 through Saturday, 12/05/2020.

Outbreak Setting Description	Ever reported	Reported in the most recent week
Food service/restaurant	199	11
Construction	130	3
Childcare	129	10
Agriculture/employer housing/produce packing	126	1
Manufacturing (non-food)	110	10
Other	102	3
Grocery	94	14
Retail	92	13
K-12 school	88	12
Manufacturing (food and food-related)	80	3
Shelter/homeless service	77	64
Professional services/office-based (business, IT, finance, legal)	55	4
Place of worship	55	1
Private event	54	2
College/university	54	0
Transportation/shipping/delivery	47	2
Corrections	40	4
Leisure/hospitality/recreation	33	4
Warehousing	31	0
Agency, facility, etc.	30	1
Congregate housing (not employer provided)	24	1
Military	24	1
Bar/nightclub	22	0
Utilities	21	1
Facility/domestic cleaning service	16	3
Public safety	15	1
Large gathering	13	0
Personal care and service (hair, nails)	13	0
Fishing	10	0
Hotel	9	0
Real estate	9	0
College housing	8	1
Natural resources and mining	3	0
Stadium, arena, venues	3	0
Summer camp	3	0
Forestry/hunting	2	0
Office	1	0
Missing setting description	19	0
Total	*1841*	*170*

Long-Term Care Facility Outbreaks by Illness Onset Date

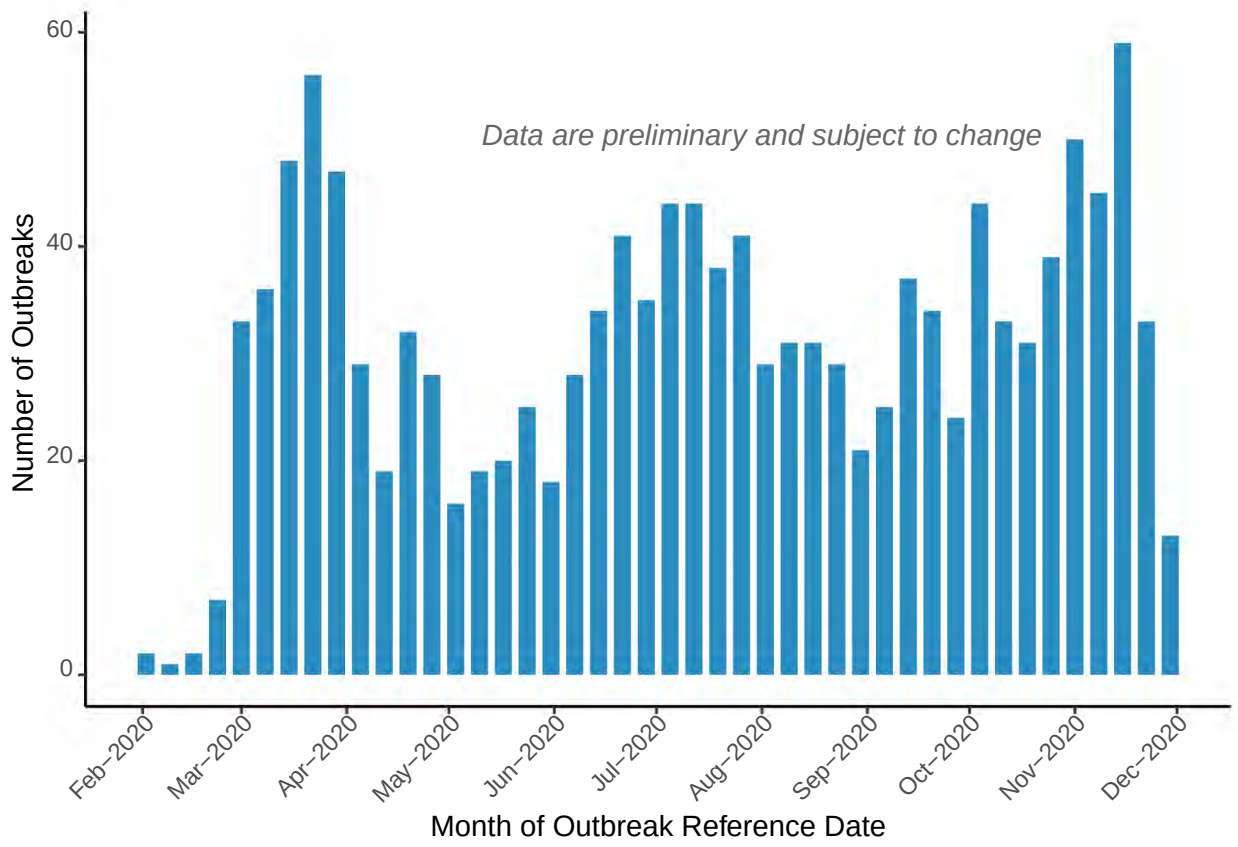
As outlined in the DOH Interim COVID-19 Outbreak Definition for Healthcare Settings, an outbreak in long-term care settings is defined by the following criteria being met:

1. One resident or healthcare worker with laboratory positive COVID-19 (PCR or antigen test), OR
2. Two or more residents or healthcare workers with new-onset respiratory symptoms consistent with COVID-19 within 72 hours of each other.

A total of 1351 long-term care (LTC) facility COVID-19 outbreaks have been reported in settings including nursing homes, assisted living facilities, and adult family homes as of 12/07/2020. This epidemiologic curve shows the number of LTC outbreaks over time by approximate onset date of the first case. Outbreaks may include residents as well as employees and visitors. Not all cases were exposed at a LTCF. Many cases visited multiple places during their exposure period, and some individuals may have visited a LTCF after disease onset. Facilities with multiple licenses for different acuity levels have been assigned to the higher acuity license.

Previously outbreaks were only reported once per facility. The Washington State Department of Health is working to provide more context to LTC outbreaks and report outbreak counts that reflect multiple outbreaks within the same facilities at least 28 days apart.

Number of Long-term Care outbreaks reported by approximate week of initial symptom onset through 12/07/2020



Non-Long-Term Care (Non-LTC) Healthcare Setting Outbreaks

As outlined in the DOH Interim COVID-19 Outbreak Definition for Healthcare Settings, an outbreak in a residential healthcare setting (e.g., residential treatment center) is defined by the following criteria being met:

1. One patient or healthcare worker with laboratory positive COVID-19 (PCR or antigen test), OR
2. Two or more patients or healthcare workers with new-onset respiratory symptoms consistent with COVID-19 within 72 hours of each other.

As outlined in the DOH Interim COVID-19 Outbreak Definition for Healthcare Settings, an outbreak in an outpatient healthcare setting (e.g., dental clinic) is defined by the following criteria being met:

1. Two or more cases of laboratory positive COVID-19 in patients or staff with epi-linkage who do not share a household, and are not listed as a close contact of each other outside of the workplace during standard case investigation or contact tracing

A total of 530 non-LTC healthcare-associated COVID-19 outbreaks have been reported in settings including hospitals, outpatient settings, behavioral health facilities, supported living facilities, home healthcare, dialysis centers, and independent senior living facilities* as of 12/05/2020. The figure below shows the number of non-LTC healthcare outbreaks over time by week of approximate onset date of the first reported case. These outbreaks may include employees and patients, and we currently have limited ability to distinguish among them. Outbreak reporting prior to June 1st is less complete than more recent weeks.

*Senior living is included in the healthcare category because these apartment-like settings are often co-located with long-term care facilities.

Number of Non-LTC Healthcare Associated outbreaks reported by approximate week of initial symptom onset through 12/05/2020

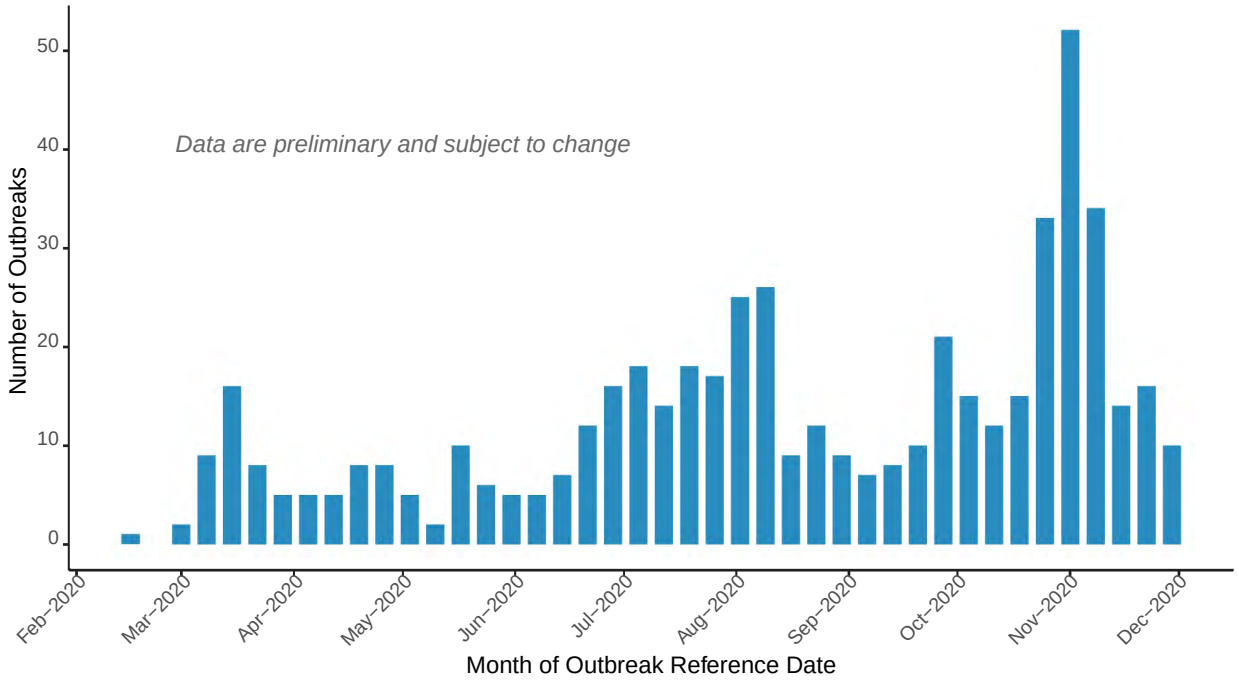


Exhibit B

Correction

ENVIRONMENTAL SCIENCES, EARTH, ATMOSPHERIC, AND PLANETARY SCIENCES

Correction for "Identifying airborne transmission as the dominant route for the spread of COVID-19," by Renyi Zhang, Yixin Li, Annie L. Zhang, Yuan Wang, and Mario J. Molina, which was first published June 11, 2020; 10.1073/pnas.2009637117 (*Proc. Natl. Acad. Sci. U.S.A.* **117**, 14857–14863).

The editors note that, due to an oversight in the proofing process, the authors' second round of edits were not incorporated into the article before publication. We apologize for the oversight and have updated the article online. In addition, the authors wish to note, "In our article the data for COVID-19 confirmed cases in Italy were retrieved from OurWorldinData.org (1), which were compiled using data from European Center for Disease Prevention and Control (ECDC) (2). However, we identified that there was a difference by one day for the total confirmed cases for Italy between the data from (1) and those reported by (2). Accordingly, the following changes have been made: (*i*) in the abstract, line 15, "78,000" has been changed to "75,000"; (*ii*) on page 14859, left column, third full paragraph, line 12, "78,000" has been changed to "75,000"; and (*iii*) Fig. 2*B* has been corrected." The article has been updated online to reflect these changes. The corrected Fig. 2 and its corresponding legend also appear below. The corrected Fig. S1 has been updated online.

Additionally, in *SI Appendix*, Fig. S2*B*, the relative humidity was misplotted. The SI has been updated accordingly.

1. M. Roser, H. Ritchie, E. Ortiz-Ospina, J. Hasell, Coronavirus Pandemic (COVID-19). https://ourworldindata.org/coronavirus. Accessed 9 May 2020.
2. European Centre for Disease Prevention and Control, Download data on the national 14-day notification rate of new COVID-19 cases and deaths. https://www.ecdc.europa.eu/en/publications-data/data-national-14-day-notification-rate-covid-19. Accessed 14 June 2020.

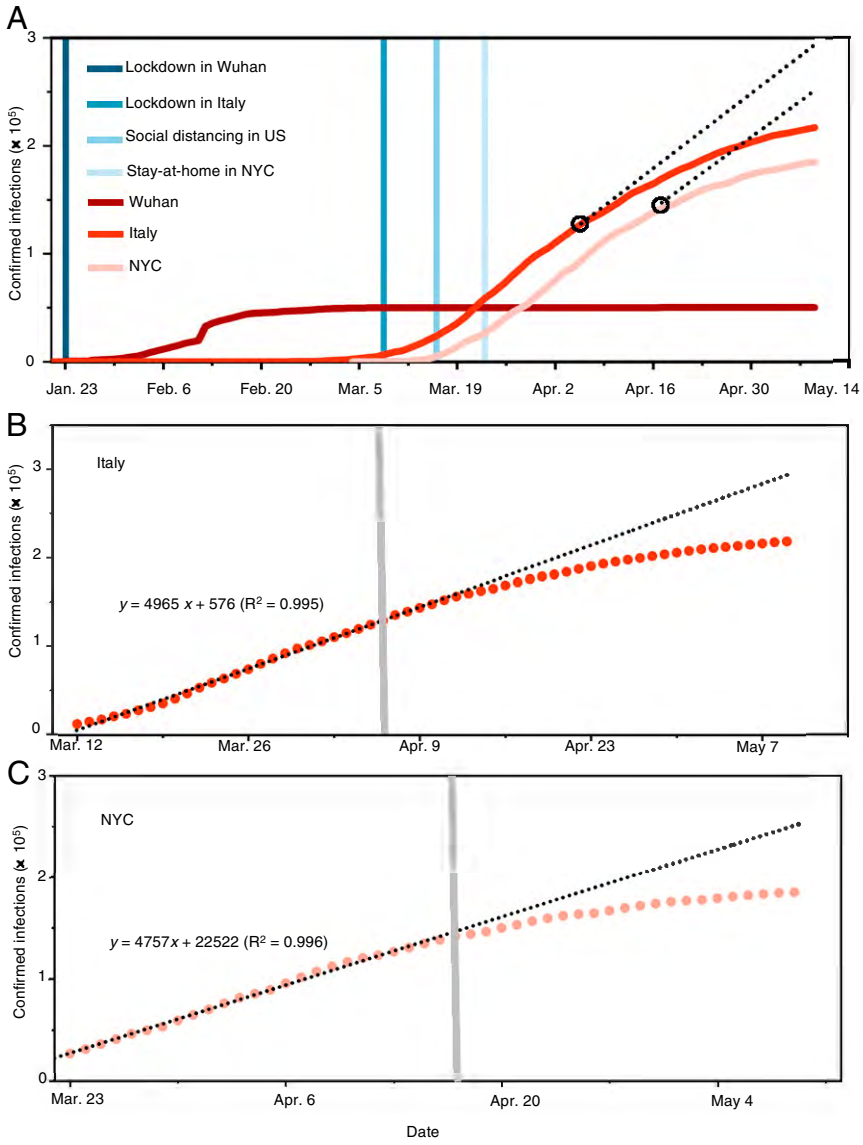


Fig. 2. The evolving epicenter from Wuhan, to Italy, to NYC. (*A*) Comparison of the trends and mitigation measures between Wuhan, Italy, and NYC in 2020. The vertical lines mark the date for implementing mitigation measures. The two black circles label the dates when face covering was implemented: April 6 in northern Italy and April 17 in NYC. The black dashed lines represent the projection without face covering based on linear regression of 26-d data prior to implementing this measure. (*B*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in Italy. The shaded vertical line denotes the date when face covering was implemented on April 6 in northern Italy. (*C*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in NYC. The shaded vertical line denotes the date when face covering was implemented on April 17 in NYC. In *B* and *C*, the circles are reported values, and the dotted line represents fitting and projection of the confirmed infections before and after face-covering, respectively.

First published October 5, 2020.

www.pnas.org/cgi/doi/10.1073/pnas.2018637117

Identifying airborne transmission as the dominant route for the spread of COVID-19

Renyi Zhang[a,b,1], Yixin Li[b], Annie L. Zhang[c], Yuan Wang[d], and Mario J. Molina[e,1]

[a]Department of Atmospheric Sciences, Texas A&M University, College Station, TX 77843; [b]Department of Chemistry, Texas A&M University, College Station, TX 77843; [c]Department of Chemistry, College of Natural Sciences, The University of Texas at Austin, Austin, TX 78712; [d]Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125; and [e]Department of Chemistry and Biochemistry, University of California San Diego, La Jolla, CA 92093

Contributed by Mario J. Molina, May 16, 2020 (sent for review May 14, 2020; reviewed by Manish Shrivastava and Tong Zhu)

Various mitigation measures have been implemented to fight the coronavirus disease 2019 (COVID-19) pandemic, including widely adopted social distancing and mandated face covering. However, assessing the effectiveness of those intervention practices hinges on the understanding of virus transmission, which remains uncertain. Here we show that airborne transmission is highly virulent and represents the dominant route to spread the disease. By analyzing the trend and mitigation measures in Wuhan, China, Italy, and New York City, from January 23 to May 9, 2020, we illustrate that the impacts of mitigation measures are discernable from the trends of the pandemic. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the pandemic trends in the three epicenters. This protective measure alone significantly reduced the number of infections, that is, by over 75,000 in Italy from April 6 to May 9 and over 66,000 in New York City from April 17 to May 9. Other mitigation measures, such as social distancing implemented in the United States, are insufficient by themselves in protecting the public. We conclude that wearing of face masks in public corresponds to the most effective means to prevent interhuman transmission, and this inexpensive practice, in conjunction with simultaneous social distancing, quarantine, and contact tracing, represents the most likely fighting opportunity to stop the COVID-19 pandemic. Our work also highlights the fact that sound science is essential in decision-making for the current and future public health pandemics.

COVID-19 | virus | aerosol | public health | pandemic

The novel coronavirus outbreak, coronavirus disease 2019 (COVID-19), which was declared a pandemic by the World Health Organization (WHO) on March 11, 2020, has infected over 4 million people and caused nearly 300,000 fatalities over 188 countries (1). Intensive effort is ongoing worldwide to establish effective treatments and develop a vaccine for the disease. The novel coronavirus, named as severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), belongs to the family of the pathogen that is responsible for respiratory illness linked to the 2002–2003 outbreak (SARS-CoV-1) (2). The enveloped virus contains a positive-sense single-stranded RNA genome and a nucleocapsid of helical symmetry of ∼120 nm. There exist several plausible pathways for viruses to be transmitted from person to person. Human atomization of virus-bearing particles occurs from coughing/sneezing and even from normal breathing/talking by an infected person (3–6). These mechanisms of viral shedding produce large droplets and small aerosols (3), which are conventionally delineated at a size of 5 μm to characterize their distinct dispersion efficiencies and residence times in air as well as the deposition patterns along the human respiratory tract (3, 7). Virus transmission occurs via direct (deposited on persons) or indirect (deposited on objects) contact and airborne (droplets and aerosols) routes (3). Large droplets readily settle out of air to cause person/object contamination; in contrast, aerosols are efficiently dispersed in air. While transmission via direct or indirect contact occurs in a short range, airborne transmission via aerosols can occur over an extended distance and time. Inhaled virus-bearing aerosols deposit directly along the human respiratory tract.

Previous experimental and observational studies on interhuman transmission have indicated a significant role of aerosols in the transmission of many respiratory viruses, including influenza virus, SARS-CoV-1, and Middle East Respiratory Syndrome coronavirus (MERS-CoV) (8–11). For example, airborne coronavirus MERS-CoV exhibited strong capability of surviving, with about 64% of microorganisms remaining infectious 60 min after atomization at 25 °C and 79% relative humidity (RH) (9). On the other hand, rapid virus decay occurred, with only 5% survival over a 60-min procedure at 38 °C and 24% RH, indicative of inactivation. Recent experimental studies have examined the stability of SARS-CoV-2, showing that the virus remains infectious in aerosols for hours (12) and on surfaces up to days (12, 13).

Several parameters likely influence the microorganism survival and delivery in air, including temperature, humidity, microbial resistance to external physical and biological stresses, and solar ultraviolet (UV) radiation (8). Transmission and infectivity of airborne viruses are also dependent on the size and number concentration of inhaled aerosols, which regulate the amount (dose) and pattern for respiratory deposition. With typical nasal breathing (i.e., at a velocity of ∼1 $m{\cdot}s^{-1}$) (4), inhalation of airborne viruses leads to direct and continuous deposition into the human respiratory tract. In particular, fine aerosols (i.e., particulate

Significance

We have elucidated the transmission pathways of coronavirus disease 2019 (COVID-19) by analyzing the trend and mitigation measures in the three epicenters. Our results show that the airborne transmission route is highly virulent and dominant for the spread of COVID-19. The mitigation measures are discernable from the trends of the pandemic. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the trends of the pandemic. This protective measure significantly reduces the number of infections. Other mitigation measures, such as social distancing implemented in the United States, are insufficient by themselves in protecting the public. Our work also highlights the necessity that sound science is essential in decision-making for the current and future public health pandemics.

Author contributions: R.Z. designed research; R.Z., Y.L., and Y.W. performed research; R.Z., Y.L., Y.W., and M.J.M. analyzed data; and R.Z., A.L.Z., and M.J.M. wrote the paper.

Reviewers: M.S., Pacific Northwest National Laboratory; and T.Z., Peking University.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: renyi-zhang@tamu.edu or mjmolina@ucsd.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2009637117/-/DCSupplemental.

First published June 11, 2020.

matter smaller than 2.5 μm, or $PM_{2.5}$) penetrate deeply into the respiratory tract and even reach other vital organs (14, 15). In addition, viral shedding is dependent on the stages of infection and varies between symptomatic and asymptomatic carriers. A recent finding (16) showed that the highest viral load in the upper respiratory tract occurs at the symptom onset, suggesting the peak of infectiousness on or before the symptom onset and substantial asymptomatic transmission for SARS-CoV-2.

The COVID-19 outbreak is significantly more pronounced than that of the 2002/2003 SARS, and the disease continues to spread at an alarming rate worldwide, despite extreme measures taken by many countries to constrain the pandemic (1). The enormous scope and magnitude of the COVID-19 outbreak reflect not only a highly contagious nature but also exceedingly efficient transmission for SARS-CoV-2. Currently, the mechanisms to spread the virus remain uncertain (17), particularly considering the relative contribution of the contact vs. airborne transmission routes to this global pandemic. Available epidemiological (1) and experimental (12, 18) evidence, however, implicates airborne transmission of SARS-CoV-2 via aerosols as a potential route for the spreading of the disease.

Distinct Pandemic Trends in the Three Epicenters

To gain insight into the mechanism of the virus transmission routes and assess the effectiveness of mitigation measures, we analyzed the trend of the pandemic worldwide from January 23 to May 9, 2020 (Fig. 1). The COVID-19 outbreak initially emerged during December 2019 in Wuhan, China (1). The numbers of confirmed infections and fatalities in China dominated the global trend during January and February 2020 (Fig. 1 *A* and *B*), but the increases in the newly confirmed cases and fatalities in China have exhibited sharp declines since February (Fig. 1*B*). In contrast to the curve flattening in China, those numbers in other countries have increased sharply since the beginning of March. The epicenter shifted from Wuhan to Italy in early March and to New York City (NYC) in early April. By April 30, the numbers of confirmed COVID-19 cases and deaths, respectively, reached over 200,000 and 27,000 in Italy and over 1,000,000 and 52,000 in the United States, compared to about 84,000 and 4,600 in China (Fig. 1*B*). Notably, the curves in Italy exhibit a slowing trend since mid-April, while the numbers in the world and the United States continue to increase. Remarkably, the recent trends in the numbers of infections and fatalities in the world and in the United States exhibit striking linearity since the beginning of April (Fig. 1*C*).

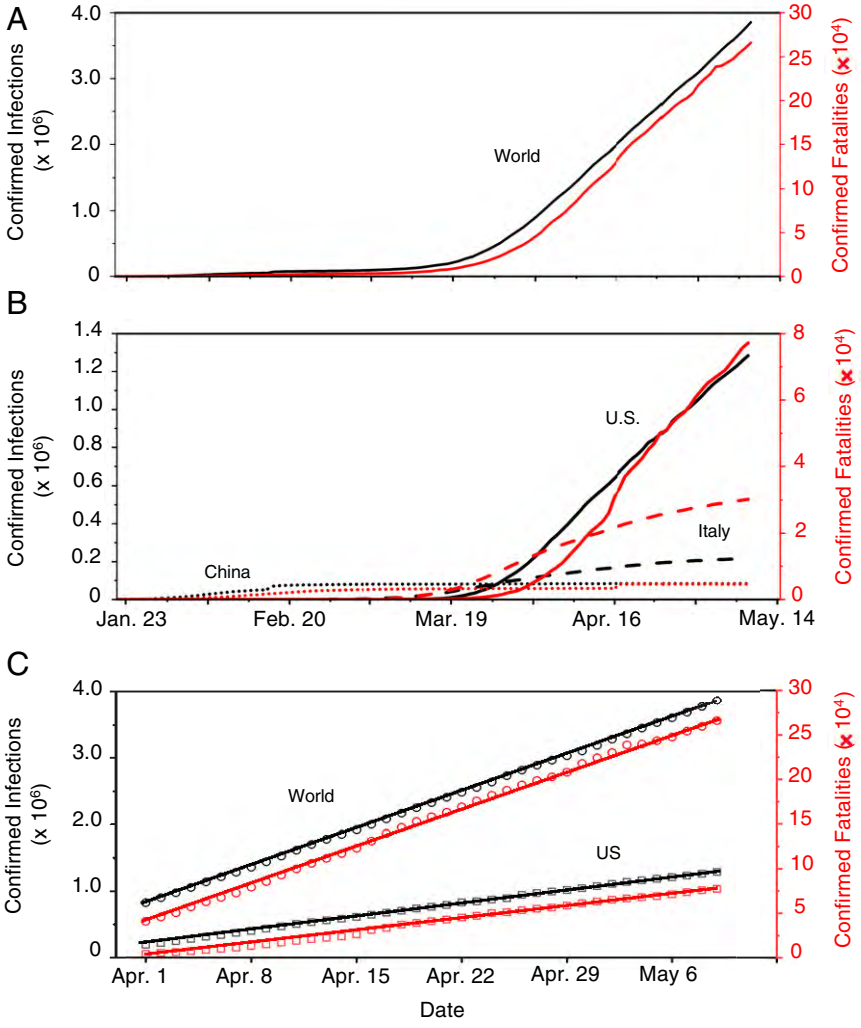


Fig. 1. Distinct global trends of the COVID-19 pandemic. (*A*) Confirmed infections and fatalities worldwide. (*B*) Comparison of the confirmed infections and fatalities between China, Italy, and United States. (*C*) Linear regression of the confirmed infections and fatalities worldwide and in United States from April 1 to May 9, 2020; the linear regression is, respectively, $y = 79{,}398x + 810{,}167$ ($R^2 = 0.999$) for infections and $y = 6{,}075x + 39{,}409$ ($R^2 = 0.998$) for fatalities worldwide and $y = 28{,}971x + 201{,}187$ ($R^2 = 0.999$) for infections and $y = 2{,}059x + 243$ ($R^2 = 0.995$) for fatalities in the United States. The left axis and black color correspond to the numbers of confirmed infections, and the right axis and red color represent the confirmed fatalities.

We interpreted the differences in the pandemic trends by considering the mitigation measures implemented worldwide. The curve flattening in China can be attributed to extensive testing, quarantine, and contact tracing; other aggressive measures implemented in China include lockdown of all cities and rural areas in the whole country, isolation of residents having close contact with infected people, and mandated wearing of face masks in public. However, the effectiveness of those mitigation measures has yet to be rigorously evaluated. Differentiation of the effects of those mitigation measures in China is challenging (19), since the implementation occurred almost simultaneously in January 2020. While similar quarantine, isolation, and city lockdown measures were also implemented on March 9 in Italy after the country became the second epicenter, the curve of infections has yet to show complete flattening. In the United States, guidelines for social distancing, quarantine, and isolation were issued by the federal government on March 16, and stay-at-home orders were implemented by many state and local governments starting, for example, between March 19 and April 3 and on March 22 in NYC. The social distancing measures implemented in the United States include staying at least 6 feet (~2 m) away from other people, no gathering in groups, staying out of crowded places, and avoiding mass gatherings (20). Obviously, the continuous rise in the US infected numbers casts doubt on the effectiveness of those preventive measures alone (Fig. 1 *B* and *C*).

In contrast to China, wearing of face masks was not mandated and was unpopular in most of the western world during the early outbreak of the pandemic. Advice on the use of face masks was not issued until April 6, 2020 by the WHO (1), claiming that it is important only to prevent infected persons from viral transmission by filtering out droplets but that it is unimportant to prevent uninfected persons from breathing virus-bearing aerosols. The regions heavily plagued by COVID-19 in northern Italy, such as Lombard, ordered face covering in public starting on April 6, and the Italian authorities required nationwide mandatory use of face masks on May 4. All New Yorkers were mandated to use face covering in public starting on April 17, when social distancing was not possible. With measures implemented in the United States seemingly comparable to those in China, social distancing, quarantine, and isolation exhibited little impact on stopping the spreading of the disease in the United States, as reflected by the linearity from April 1 to May 9 (Fig. 1*C*). It is possible, however, that these measures alter the slope of the infection curve, that is, by reducing the rate of infections during the early stage of the pandemic (Fig. 1). Notably, the recommended physical separation for social distancing is beneficial to prevent direct contact transmission but is insufficient (without face masks) to protect inhalation of virus-bearing aerosols (or even small droplets at intermediate proximity), owing to rapid air mixing (7).

Understanding the Impacts of Face Covering

Compared to the simultaneous implementation of measures in China, intervention measures were successively implemented in the western world (Fig. 2*A*), providing an opportunity for assessing their relative effectiveness. We quantified the effects of face covering by projecting the number of infections based on the data prior to implementing the use of face masks in Italy on April 6 and NYC on April 17 (Fig. 2*A*; see *Methods*). Such projections are reasonable considering the excellent linear correlation for the data prior to the onset of mandated face covering (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1). Our analysis indicates that face covering reduced the number of infections by over 75,000 in Italy from April 6 to May 9 and by over 66,000 in NYC from April 17 to May 9. In addition, varying the correlation from 15 d to 30 d prior to the onset of the implementation reveals little difference in the projection for both places, because of the high correlation coefficients (*SI Appendix*, Fig. S1). Notably, the trends of the infection curves in Italy and NYC contrast to those in the world and in the United States (Fig. 1*C*), which show little deviation from the linearity due to the nonimplementation of face-covering measures globally and nationally, respectively. The inability of social distancing, quarantine, and isolation alone to curb the spread of COVID-19 is also evident from the linearity of the infection curve prior to the onset of the face-covering rule in Italy on April 6 and in NYC on April 17 (Fig. 2 *B* and *C*). Hence, the difference made by implementing face covering significantly shapes the pandemic trends worldwide.

We further compared the numbers of daily new cases between NYC and the United States (excluding the data in New York State) from March 1 to May 9 (Fig. 3). The daily numbers of newly confirmed infections in NYC and the United States show a sharp increase in late March and early April. There exists a slower increase in the number after implementation of the stay-at-home order (about 14 d in NYC and shortly after April 3 in the United States), which is attributable to the impacts of this measure. After April 3, the only difference in the regulatory measures between NYC and the United States lies in face covering on April 17 in NYC. We applied linear regression to the data between April 17 and May 9 in NYC and between April 5 and May 9 in the United States. While the daily numbers of newly confirmed infections fluctuate considerably, the slope of the regression unambiguously reflects the trend in both data. The daily new infection in NYC decreases with a slope of 106 cases per day after April 17, corresponding to a decreasing rate of ~3% per day (relative to April 17). For comparison, the daily new infections in the United States (excluding New York State) increase, with a slope of 70 cases per day after April 4, corresponding to an increasing trend of ~0.3% per day (relative to April 5). Hence, the decreasing rate in the daily new infections in NYC with mandated face covering is in sharp contrast to that in the United States with only social-distancing and stay-at-home measures, further confirming the importance of face covering in intervening the virus transmission.

Dominant Airborne Transmission

We further elucidated the contribution of airborne transmission to the COVID-19 outbreak by comparing the trends and mitigation measures during the pandemic worldwide and by considering the virus transmission routes (Fig. 4). Face covering prevents both airborne transmission by blocking atomization and inhalation of virus-bearing aerosols and contact transmission by blocking viral shedding of droplets. On the other hand, social distancing, quarantine, and isolation, in conjunction with hand sanitizing, minimize contact (direct and indirect) transmission but do not protect against airborne transmission. With social distancing, quarantine, and isolation in place worldwide and in the United States since the beginning of April, airborne transmission represents the only viable route for spreading the disease, when mandated face covering is not implemented. Similarly, airborne transmission also contributes dominantly to the linear increase in the infection prior to the onset of mandated face covering in Italy and NYC (Fig. 2 *B* and *C* and *SI Appendix*, Fig. S1). Hence, the unique function of face covering to block atomization and inhalation of virus-bearing aerosols accounts for the significantly reduced infections in China, Italy, and NYC (Figs. 1–3), indicating that airborne transmission of COVID-19 represents the dominant route for infection.

Recent measurements identified SARS-Cov-2 RNA on aerosols in Wuhan's hospitals (18) and outdoor in northern Italy (21), unraveling the likelihood of indoor and outdoor airborne transmission. Within an enclosed environment, virus-bearing aerosols from human atomization are readily accumulated, and elevated levels of airborne viruses facilitate transmission from person to person. Transmission of airborne viruses in open air is subject to

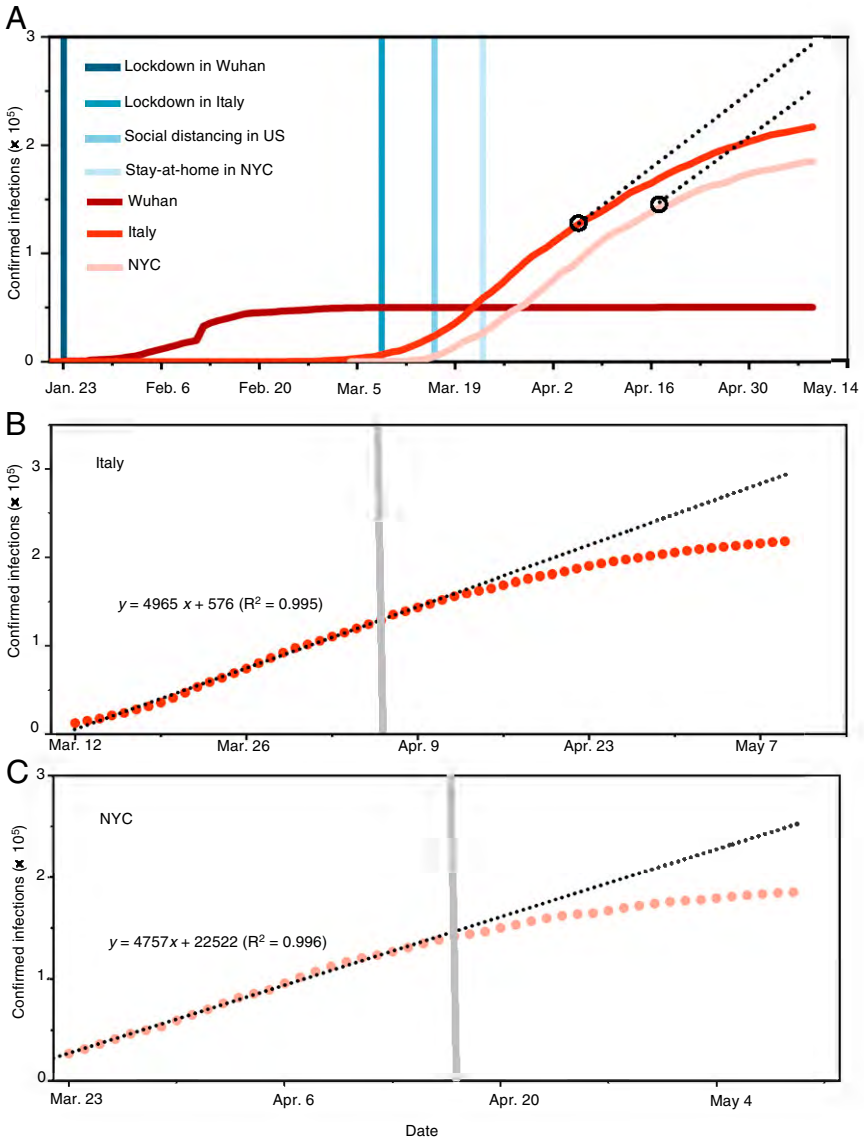


Fig. 2. The evolving epicenter from Wuhan, to Italy, to NYC. (*A*) Comparison of the trends and mitigation measures between Wuhan, Italy, and NYC in 2020. The vertical lines mark the date for implementing mitigation measures. The two black circles label the dates when face covering was implemented: April 6 in northern Italy and April 17 in NYC. The black dashed lines represent the projection without face covering based on linear regression of 26-d data prior to implementing this measure. (*B*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in Italy. The shaded vertical line denotes the date when face covering was implemented on April 6 in northern Italy. (*C*) Linear regression of the number of confirmed infections for 26-d data prior to implementing face covering in NYC. The shaded vertical line denotes the date when face covering was implemented on April 17 in NYC. In *B* and *C*, the circles are reported values, and the dotted line represents fitting and projection of the confirmed infections before and after face-covering, respectively.

dilution, although virus accumulation still occurs due to stagnation under polluted urban conditions (7, 22). Removal of virus-bearing particles from human atomization via deposition is strongly size dependent, with the settling velocities ranging from 2.8×10^{-5} m·s^{-1} to 1.4×10^{-3} m·s^{-1} for the sizes of 1 and 10 μm, respectively (7). For comparison, typical wind velocity is about 1 m·s^{-1} to 3 m·s^{-1} indoors (23) and is ∼1 m·s^{-1} horizontally and 0.1 m·s^{-1} vertically in stable air (7, 22). Under those indoor and outdoor conditions, the residence time of virus-bearing aerosols reaches hours, due to air mixing (7).

We also examined ambient conditions relevant to the outbreaks in Wuhan, Italy, and NYC. The initial outbreak of COVID-19 in Wuhan coincided with the winter haze season in China (7, 22), during which high levels of $PM_{2.5}$ were prevalent in air (*SI Appendix*, Figs. S2 and S3). On the other hand, the daily average $PM_{2.5}$ concentrations were much lower during the outbreaks in Rome, Italy, and in NYC (*SI Appendix*, Fig. S2). The airborne transmission pathways (i.e., indoor or outdoor) as well as the effects of ambient $PM_{2.5}$ levels on virus transmission may be variable among urban cities. For example, the winter haze conditions in China likely exacerbated outdoor virus spreading (24, 25), because of low UV radiation, air stagnation (lacking ventilation on the city scale), and low temperature (7, 22). Also, there may exist a synergetic effect of simultaneous exposure to the virus and $PM_{2.5}$ to enhance the infectivity, severity, and fatalities of the disease (14, 26). In addition, nascent virus-bearing aerosols produced from human atomization likely undergo transformation in air, including coagulation with ambient preexisting PM and/or growth on a time scale of a few hours in typical urban air (27–29). Such transformation, as recently documented on coarse PM in Italy (21), may mitigate

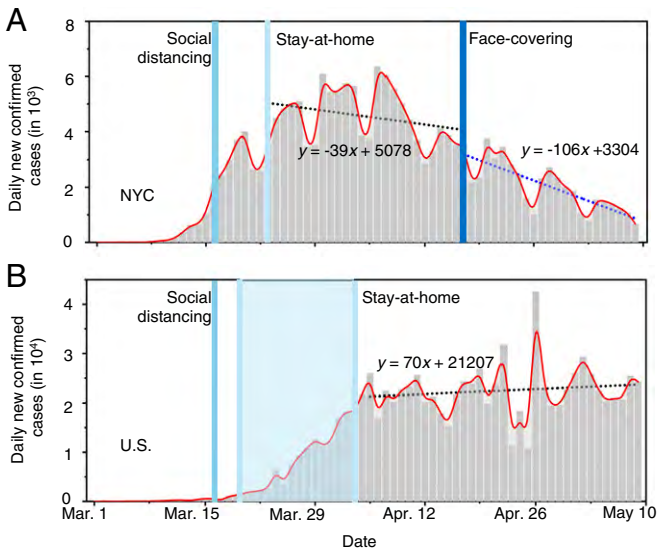


Fig. 3. Contrasting the trends of new infections between NYC and the United States. Daily new confirmed infections in (*A*) NYC and (*B*) the United States. The dotted lines represent linear fitting to the data between April 17 and May 9 in NYC and between April 4 and May 9 in the United States. In *B*, the number in New York State was subtracted from that in the United States. The vertical lines label the dates for social distancing, stay-at-home orders (the shaded area for the range of implementing dates for different states), and mandated face-covering.

virus inactivation (9, 12), by providing a medium to preserve its biological properties and elongating its lifetimes. However, key questions remain concerning transformation and transmission of virus-bearing aerosols from human atomization in air. Specifically, what are the impacts of transformation of human-atomized aerosols on viral surviving and infectivity in air?

While the humidity effect on viral surviving is uncertain (3, 9), the conditions during the outbreaks in Wuhan, Rome, and NYC correspond to high RH yet low absolute humidity because of low temperature (*SI Appendix*, Fig. S3). Early experimental work (9) showed remarkable survival for the analogous coronavirus MERS-CoV at the RH level characteristic of the COVID-19 outbreaks in Wuhan, Rome, and NYC. For comparison, indoor temperature and RH typically range from 21 °C to 27 °C and 20 to 70%, respectively (23).

Of particular importance are the considerations that render airborne SARS-CoV-2 the most efficient among all transmission routes. Even with normal nasal breathing, inhalation of virus-bearing aerosols results in deep and continuous deposition into the human respiratory tract, and this transmission route typically requires a low dose (8). Also, airborne viruses have great mobility and sufficiently long surviving time for dispersion (9, 12), and residents situated in densely populated environments are highly vulnerable. In addition, nascent micrometer-size aerosols produced from coughing/sneezing of infected people have the potential of containing many viruses, particularly for asymptomatic carriers (16).

Future research is critically needed to assess the transmission, transformation, and dispersion of virus-bearing aerosols from human atomization under different environmental conditions, as well as the related impacts on virus infectivity. It is equally important to understand human atomization of airborne viruses: What are the number and size distributions of nascent aerosols as well as the viral load per particle from coughing/sneezing? It is also imperative to evaluate human inhalation of airborne viruses: How are aerosols deposited along the respiratory tract, and what is the minimum dose of airborne viruses required for infection? It is also important to evaluate the performance of face masks to quantify the efficiency to filtrate airborne viruses relevant to human atomization and inhalation. Elucidation of these mechanisms requires an interdisciplinary effort.

A Policy Perspective

The governments' responses to the COVID-19 pandemic have so far differed significantly worldwide. Swift actions to the initial outbreak were undertaken in China, as reflected by nearly simultaneous implementation of various aggressive mitigation measures. On the other hand, the response to the pandemic was generally slow in the western world, and implementation of the intervention measures occurred only consecutively. Clearly, the responsiveness of the mitigation measures governed the evolution, scope, and magnitude of the pandemic globally (Figs. 1 and 2).

Curbing the COVID-19 relies not only on decisive and sweeping actions but also, critically, on the scientific understanding of the virus transmission routes, which determines the effectiveness of the mitigation measures (Fig. 5). In the United States, social distancing and stay-at-home measures, in conjunction with hand sanitizing (Fig. 5, path A), were implemented during the early stage of the pandemic (20). These measures minimized short-range contact transmission but did not prevent long-range airborne transmission, responsible for the inefficient containing of the pandemic in the United States (Figs. 1 and 3). Mandated face covering, such as those implemented in China, Italy, and NYC, effectively prevented airborne transmission by blocking atomization and inhalation of virus-bearing aerosols and contact transmission by blocking viral shedding of droplets. While the combined face-covering and social distancing measures offered dual protection

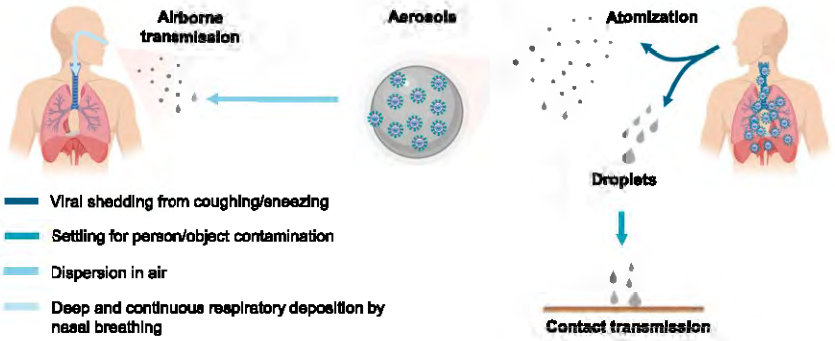


Fig. 4. Transmission of COVID-19. Human atomization of viruses arises from coughing or sneezing of an infected person, producing virus-containing droplets (>5 μm) and aerosols (<5 μm). Virus transmission from person to person occurs through direct/indirect contact and airborne aerosol/droplet routes. Large droplets mainly settle out of air to cause person/object contamination, while aerosols are efficiently dispersed in air. Direct and airborne transmissions occur in short range and extended distance/time, respectively. Inhaled airborne viruses deposit directly into the human respiration tract.

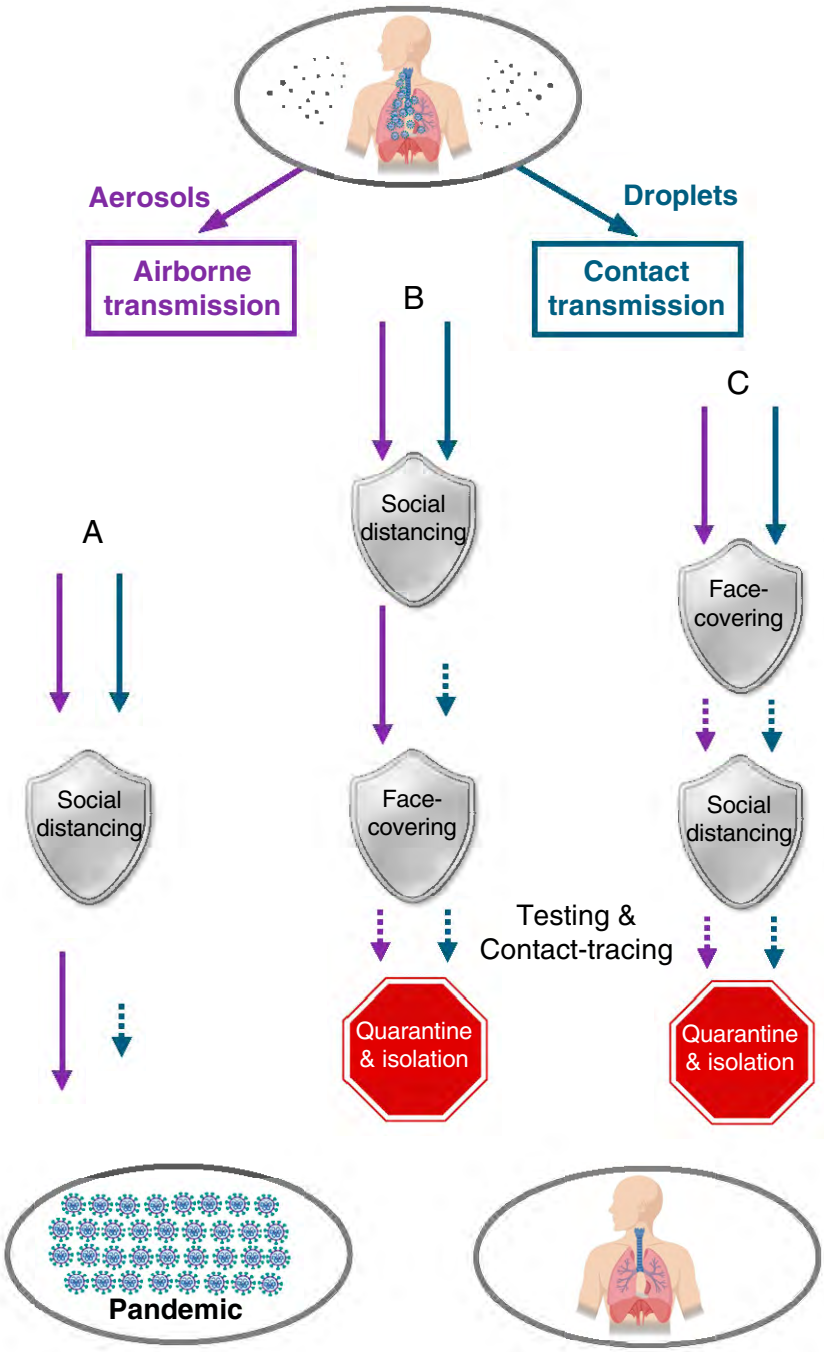


Fig. 5. Mitigation paradigm. Scenarios of virus transmission under the distancing/quarantine/isolation measure only (path A), the measures with distancing/quarantine/isolation followed by face covering (path B), and the measures with simultaneous face covering and distancing/quarantine/isolation (path C). The short-dashed arrows label possible remnants of virus transmission due to circumstances when the measure is not possible or disobeyed and/or imperfection of the measure.

against the virus transmission routes, the timing and sequence in implementing the measures also exhibited distinct outcomes during the pandemic. For example, social distancing measures, including city lockdown and stay-at-home orders, were implemented well before face covering was mandated in Italy and NYC (Fig. 5, path B), and this sequence left an extended window (28 d in Italy and 32 d in NYC) for largely uninterrupted airborne transmission to spread the disease (Figs. 2 and 3). The simultaneous implementation of face covering and social distancing (Fig. 5, path C), such as that undertaken in China, was most optimal, and this configuration, in conjunction with extensive testing and contact tracing, was responsible for the curve flattening in China (Fig. 1). Also, there likely existed remnants of virus transmission after the implementation of regulatory measures, because of circumstances when the measures were not practical or were disobeyed and/or imperfection of the measures. Such limitations, which have been emphasized by the WHO (1), spurred on controversial views on the validity of wearing face masks to prevent the virus transmission during the pandemic (30). However, it is implausible that the limitations of mitigation measures alone contributed dominantly to the global pandemic trend, as exemplified by the success in China. Our work suggests that the failure in containing the propagation of COVID-19 pandemic worldwide is largely attributed to the unrecognized importance of airborne virus transmission (1, 20).

Conclusions

The inadequate knowledge on virus transmission has inevitably hindered development of effective mitigation policies and resulted in unstoppable propagation of the COVID-19 pandemic (Figs. 1–3). In this work, we show that airborne transmission, particularly via nascent aerosols from human atomization, is highly virulent and represents the dominant route for the transmission of this disease. However, the importance of airborne transmission has not been considered in establishment of mitigation measures by government authorities (1, 20). Specifically, while the WHO and the US Centers for Disease Control and Prevention (CDC) have emphasized the prevention of contact transmission, both WHO and CDC have largely ignored the importance of the airborne transmission route (1, 20). The current mitigation measures, such as social distancing, quarantine, and isolation implemented in the United States, are insufficient by themselves in protecting the public. Our analysis reveals that the difference with and without mandated face covering represents the determinant in shaping the trends of the pandemic worldwide. We conclude that wearing of face masks in public corresponds to the most effective means to prevent interhuman transmission, and this inexpensive practice, in conjunction with extensive testing, quarantine, and contact tracing, poses the most probable fighting opportunity to stop the COVID-19 pandemic, prior to the development of a vaccine. It is also important to emphasize that sound science should be effectively communicated to policy makers and should constitute the prime foundation in decision-making amid this pandemic. Implementing policies without a scientific basis could lead to catastrophic consequences, particularly in light of attempts to reopen the economy in many countries. Clearly, integration between science and policy is crucial to formulation of effective emergency responses by policy makers and preparedness by the public for the current and future public health pandemics.

Methods

Projection of the pandemic trend without implementing face covering in Italy and NYC was performed first by establishing the linear correlation between the infection number and date. We considered the data for both 15 and 26 d prior to the onset of face covering (*SI Appendix*, Fig. S1). The slope and the reported infection number were used for the projections. The avoided infection number due the face covering was determined from the difference between the projected and reported values on May 9, 2020.

The data for accumulative confirmed infections and fatalities in Wuhan, Italy, and NYC were taken from the reports by Wuhan Municipal Health Commission (http://wjw.wuhan.gov.cn/), European CDC (https://www.ecdc.europa.eu/en), and NYC government (https://www1.nyc.gov/site/doh/covid/covid-19-data.page), respectively. The data of accumulative confirmed infections and fatalities worldwide were taken from WHO COVID-19 situation report (https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports) (1), and the numbers in China, Italy, and United States were from taken from European CDC.

Ground-based measurements of $PM_{2.5}$ and RH in Wuhan were taken from the China National Environmental Monitoring Centre (http://beijingair.sinaapp.com/). The $PM_{2.5}$ data in NYC were taken from US Environmental Protection Agency (https://www.epa.gov/outdoor-air-quality-data). The $PM_{2.5}$ data in Rome were taken were from Centro Regionale della Qualità dell'aria (http://www.arpalazio.net/main/aria/). The RH data in Rome and NYC were taken from the 6-hourly interim reanalysis of the European Centre for Medium-range Weather Forecasts (https://www.ecmwf.int/en/forecasts/datasets/reanalysis-datasets/era5).

We used spaceborne measurements of aerosol optical depth (AOD) to characterize the regional aerosol pollution during the COVID-19 outbreak (January 23 to February 10, 2020) in China. The green band AODs at 0.55 μm are available from Terra and Aqua combined Moderate Resolution Imaging Spectroradiometer Version 6 Multiangle Implementation of Atmospheric Correction (https://lpdaac.usgs.gov/products/mcd19a2v006/). The Level-2 product has daily global coverage with 1-km pixel resolution. The AOD retrieval is only available for the clear sky.

Data Availability. All data relevant to this research are available in the main text and *SI Appendix*.

ACKNOWLEDGMENTS. This work was supported by the Robert A. Welch Foundation (Grant A-1417). A.L.Z. acknowledges the support of a fellowship from the Robert A. Welch Foundation. We are grateful to Fang Zhang for the $PM_{2.5}$ data in Wuhan, China.

1. World Health Organization, Coronavirus disease (COVID-2019) situation reports. https://www.who.int/emergencies/diseases/novel-coronavirus-2019/situation-reports/. Accessed 9 May 2020.
2. A. R. Fehr, S. Perlman, Coronaviruses: An overview of their replication and pathogenesis. *Methods Mol. Biol.* **1282**, 1–23 (2015).
3. J. S. Kutter, M. I. Spronken, P. L. Fraaij, R. A. Fouchier, S. Herfst, Transmission routes of respiratory viruses among humans. *Curr. Opin. Virol.* **28**, 142–151 (2018).
4. J. W. Tang *et al.*, Airflow dynamics of human jets: Sneezing and breathing - potential sources of infectious aerosols. *PLoS One* **8**, e59970 (2013).
5. N. H. L. Leung *et al.*, Respiratory virus shedding in exhaled breath and efficacy of face masks. *Nat. Med.* **26**, 676–680 (2020).
6. V. Stadnytskyi, C. E. Bax, A. Bax, P. Anfinrud, The airborne lifetime of small speech droplets and their potential importance in SARS-CoV-2 transmission. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 11875–11877 (2020).
7. R. Zhang *et al.*, Formation of urban fine particulate matter. *Chem. Rev.* **115**, 3803–3855 (2015).
8. R. Tellier, Aerosol transmission of influenza A virus: A review of new studies. *J. R. Soc. Interface* **6** (suppl. 6), S783–S790 (2009).
9. O. V. Pyankov, S. A. Bodnev, O. G. Pyankova, I. E. Agranovski, Survival of aerosolized coronavirus in the ambient air. *J. Aerosol Sci.* **115**, 158–163 (2018).
10. M. Richard, R. A. M. Fouchier, Influenza A virus transmission via respiratory aerosols or droplets as it relates to pandemic potential. *FEMS Microbiol. Rev.* **40**, 68–85 (2016).
11. T. P. Weber, N. I. Stilianakis, Inactivation of influenza A viruses in the environment and modes of transmission: A critical review. *J. Infect.* **57**, 361–373 (2008).
12. N. van Doremalen *et al.*, Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N. Engl. J. Med.* **382**, 1564–1567 (2020).
13. A. W. H. Chin *et al.*, Stability of SARS-CoV-2 in different environmental conditions. *Lancet* **1**, E10 (2020).
14. K. A. Rychlik *et al.*, In utero ultrafine particulate matter exposure causes offspring pulmonary immunosuppression. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 3443–3448 (2019).
15. G. Wu *et al.*, Adverse organogenesis and predisposed long-term metabolic syndrome from prenatal exposure to fine particulate matter. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 11590–11595 (2019).
16. X. He *et al.*, Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat. Med.* **26**, 672–675 (2020).
17. D. Lewis, Is the coronavirus airborne? Experts can't agree. *Nature* **580**, 175 (2020).
18. Y. Liu *et al.*, Aerodynamic analysis of SARS-CoV-2 in two Wuhan hospitals. *Nature*, 10.1038/s41586-020-2271-3 (2020).
19. L. Ferretti *et al.*, Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing. *Science* **368**, eabb6936 (2020).
20. US Centers for Disease Control and Prevention, Coronavirus Disease 2019 (COVID-19) - Social distancing, quarantine, and isolation. https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/social-distancing.html. Accessed 9 May 2020.
21. L. Setti *et al*, SARS-Cov-2 RNA found on particulate matter of Bergamo in Northern Italy: First preliminary evidence. *Environ. Res.*, 10.1016/j.envres.2020.109754 (2020).0013-9351
22. Z. An *et al.*, Severe haze in northern China: A synergy of anthropogenic emissions and atmospheric processes. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 8657–8666 (2019).
23. L. A. Wallace, S. J. Emmerich, C. Howard-Reed, Continuous measurements of air change rates in an occupied house for 1 year: The effect of temperature, wind, fans, and windows. *J. Expo. Anal. Environ. Epidemiol.* **12**, 296–306 (2002).
24. Q. Ye, J. F. Fu, J. H. Mao, S. Q. Shang, Haze is a risk factor contributing to the rapid spread of respiratory syncytial virus in children. *Environ. Sci. Pollut. Res. Int.* **23**, 20178–20185 (2016).
25. Z. Gong *et al.*, Probable aerosol transmission of severe fever with thrombocytopenia syndrome virus in southeastern China. *Clin. Microbiol. Infect.* **21**, 1115–1120 (2015).
26. X. Wu *et al.*, Exposure to air pollution and COVID-19 mortality in the United States. https://projects.iq.harvard.edu/files/covid-pm/files/pm_and_covid_mortality.pdf. Accessed 9 May 2020.
27. S. Guo *et al.*, Elucidating severe urban haze formation in China. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 17373–17378 (2014).
28. F. Zhang *et al.*, An unexpected catalyst dominates formation and radiative forcing of regional haze. *Proc. Natl. Acad. Sci. U.S.A.* **117**, 3960–3966 (2020).
29. J. Peng *et al.*, Markedly enhanced absorption and direct radiative forcing of black carbon under polluted urban environments. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 4266–4271 (2016).
30. J. Howard *et al*, Face masks against COVID-19: An evidence review. https://doi.org/10.20944/preprints202004.0203.v2 (13 May 2020).

Exhibit C

Article

Mobility network models of COVID-19 explain inequities and inform reopening

https://doi.org/10.1038/s41586-020-2923-3

Received: 15 June 2020

Accepted: 21 October 2020

Published online: 10 November 2020

Check for updates

Serina Chang[1,9], Emma Pierson[1,2,9], Pang Wei Koh[1,9], Jaline Gerardin[3], Beth Redbird[4,5], David Grusky[6,7] & Jure Leskovec[1,8]✉

The coronavirus disease 2019 (COVID-19) pandemic markedly changed human mobility patterns, necessitating epidemiological models that can capture the effects of these changes in mobility on the spread of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2)[1]. Here we introduce a metapopulation susceptible–exposed–infectious–removed (SEIR) model that integrates fine-grained, dynamic mobility networks to simulate the spread of SARS-CoV-2 in ten of the largest US metropolitan areas. Our mobility networks are derived from mobile phone data and map the hourly movements of 98 million people from neighbourhoods (or census block groups) to points of interest such as restaurants and religious establishments, connecting 56,945 census block groups to 552,758 points of interest with 5.4 billion hourly edges. We show that by integrating these networks, a relatively simple SEIR model can accurately fit the real case trajectory, despite substantial changes in the behaviour of the population over time. Our model predicts that a small minority of 'superspreader' points of interest account for a large majority of the infections, and that restricting the maximum occupancy at each point of interest is more effective than uniformly reducing mobility. Our model also correctly predicts higher infection rates among disadvantaged racial and socioeconomic groups[2–8] solely as the result of differences in mobility: we find that disadvantaged groups have not been able to reduce their mobility as sharply, and that the points of interest that they visit are more crowded and are therefore associated with higher risk. By capturing who is infected at which locations, our model supports detailed analyses that can inform more-effective and equitable policy responses to COVID-19.

In response to the COVID-19 crisis, stay-at-home orders were enacted in many countries to reduce contact between individuals and slow the spread of the SARS-CoV-2[9]. Since then, public officials have continued to deliberate over when to reopen, which places are safe to return to and how much activity to allow[10]. Answering these questions requires epidemiological models that can capture the effects of changes in mobility on virus spread. In particular, findings of COVID-19 superspreader events[11–14] motivate models that can reflect the heterogeneous risks of visiting different locations, whereas well-reported disparities in infection rates among different racial and socioeconomic groups[2–8] require models that can explain the disproportionate effect of the virus on disadvantaged groups.

To address these needs, we construct fine-grained dynamic mobility networks from mobile-phone geolocation data, and use these networks to model the spread of SARS-CoV-2 within 10 of the largest metropolitan statistical areas (hereafter referred to as metro areas) in the USA. These networks map the hourly movements of 98 million people from census block groups (CBGs), which are geographical units that typically contain 600–3,000 people, to specific points of interest (POIs). As shown in Supplementary Table 1, POIs are non-residential locations that people visit such as restaurants, grocery stores and religious establishments. On top of each network, we overlay a metapopulation SEIR model that tracks the infection trajectories of each CBG as well as the POIs at which these infections are likely to have occurred. This builds on prior research that models disease spread using aggregate[15–19], historical[20–22] or synthetic mobility data[23–25]; separately, other studies have analysed mobility data in the context of COVID-19, but without an underlying model of disease spread[26–30].

Combining our epidemiological model with these mobility networks allows us to not only accurately fit observed case counts, but also to conduct detailed analyses that can inform more-effective and equitable policy responses to COVID-19. By capturing information about individual POIs (for example, the hourly number of visitors and median visit duration), our model can estimate the effects of specific reopening strategies, such as only reopening certain POI categories or restricting the maximum occupancy at each POI. By modelling movement from CBGs, our model can identify at-risk populations and correctly predict, solely from mobility patterns, that disadvantaged racial and

[1]Department of Computer Science, Stanford University, Stanford, CA, USA. [2]Microsoft Research, Cambridge, MA, USA. [3]Department of Preventive Medicine, Northwestern University, Chicago, IL, USA. [4]Department of Sociology, Northwestern University, Evanston, IL, USA. [5]Institute for Policy Research, Northwestern University, Evanston, IL, USA. [6]Department of Sociology, Stanford University, Stanford, CA, USA. [7]Center on Poverty and Inequality, Stanford University, Stanford, CA, USA. [8]Chan Zuckerberg Biohub, San Francisco, CA, USA. [9]These authors contributed equally: Serina Chang, Emma Pierson, Pang Wei Koh. ✉e-mail: jure@cs.stanford.edu

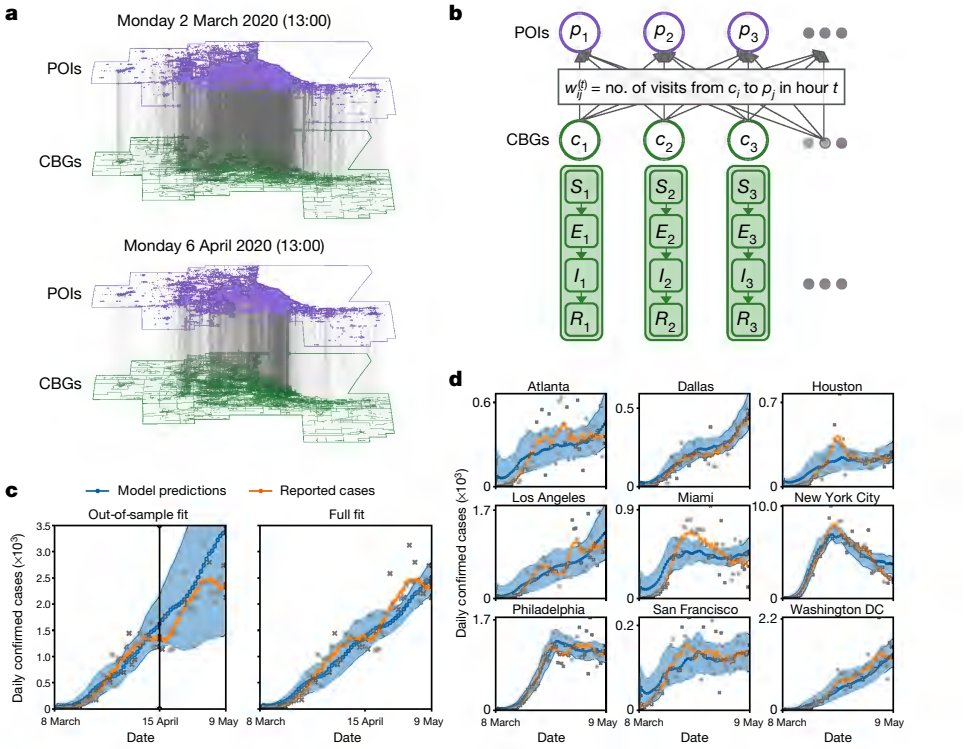


Fig. 1 | Model description and fit. a, The mobility network captures hourly visits from each CBG to each POI. The vertical lines indicate that most visits are between nearby POIs and CBGs. Visits dropped markedly from March to April, as indicated by the lower density of grey lines. Mobility networks in the Chicago metro area are shown for 13:00 on two Mondays, 2 March 2020 (top) and 6 April 2020 (bottom). **b**, We overlaid a disease-spread model on the mobility network, with each CBG having its own set of SEIR compartments. New infections occur at both POIs and CBGs, with the mobility network governing how subpopulations from different CBGs interact as they visit POIs. **c**, Left, to test the out-of-sample prediction, we calibrated the model on data before 15 April 2020 (vertical black line). Even though its parameters remain fixed over time, the model accurately predicts the case trajectory in the Chicago metro area after 15 April using the mobility data (r.m.s.e. on daily cases = 406 for dates ranging from 15 April to 9 May). Right, model fit was further improved when we calibrated the model on the full range of data (r.m.s.e. on daily cases = 387 for the dates ranging from 15 April to 9 May). **d**, We fitted separate models to 10 of the largest US metro areas, modelling a total population of 98 million people; here, we show full model fits, as in **c** (right). In **c** and **d**, the blue line represents the model predictions and the grey crosses represent the number of daily reported cases; as the numbers of reported cases tend to have great variability, we also show the smoothed weekly average (orange line). Shaded regions denote the 2.5th and 97.5th percentiles across parameter sets and stochastic realizations. Across metro areas, we sample 97 parameter sets, with 30 stochastic realizations each (n = 2,910); see Supplementary Table 6 for the number of sets per metro area.

socioeconomic groups face higher rates of infection. Our model thus enables the analysis of urgent health disparities; we use it to highlight two mobility-related mechanisms that drive these disparities and to evaluate the disparate effect of reopening on disadvantaged groups.

Mobility network model

We use data from SafeGraph, a company that aggregates anonymized location data from mobile applications, to study mobility patterns from 1 March to 2 May 2020. For each metro area, we represent the movement of individuals between CBGs and POIs as a bipartite network with time-varying edges, in which the weight of an edge between a CBG and POI represents the number of visitors from that CBG to that POI during a given hour (Fig. 1a). SafeGraph also provides the area in square feet of each POI, as well as its category in the North American industry classification system (for example, fitness centre or full-service restaurant) and median visit duration in minutes. We validated the SafeGraph mobility data by comparing the dataset to Google mobility data (Supplementary Fig. 1 and Supplementary Tables 2, 3) and used iterative proportional fitting[31] to derive POI–CBG networks from the raw SafeGraph data. Overall, these networks comprise 5.4 billion hourly edges between 56,945 CBGs and 552,758 POIs (Extended Data Table 1).

We overlay a SEIR model on each mobility network[15,20], in which each CBG maintains its own susceptible (S), exposed (E), infectious (I) and removed (R) states (Fig. 1b). New infections occur at both POIs and CBGs, with the mobility network governing how subpopulations from different CBGs interact as they visit POIs. We use the area, median visit duration and time-varying density of infectious individuals for each POI to determine the hourly infection rate of that POI. The model has only three free parameters that scale: (1) transmission rates at POIs, (2) transmission rates at CBGs and (3) the initial proportion of exposed individuals (Extended Data Table 2); all three parameters remain constant over time. We calibrate a separate model for each metro area using the confirmed case counts from *The New York Times* by minimizing the root mean square error (r.m.s.e.) to daily incident cases[32]. Our model accurately fits observed daily case counts in all 10 metro areas from 8 March to 9 May 2020 (Fig. 1c, d). In addition, when calibrated on only the case counts up to 14 April, the model predicts case counts reasonably well on the held-out time period of 15 April–9 May 2020 (Fig. 1c and Extended Data Fig. 1a). Our key technical finding is that the dynamic mobility network allows even our relatively simple SEIR model with just three static parameters to accurately fit observed cases, despite changing policies and behaviours during that period.

Mobility reduction and reopening plans

We can estimate the impact of mobility-related policies by constructing a hypothetical mobility network that reflects the expected effects of

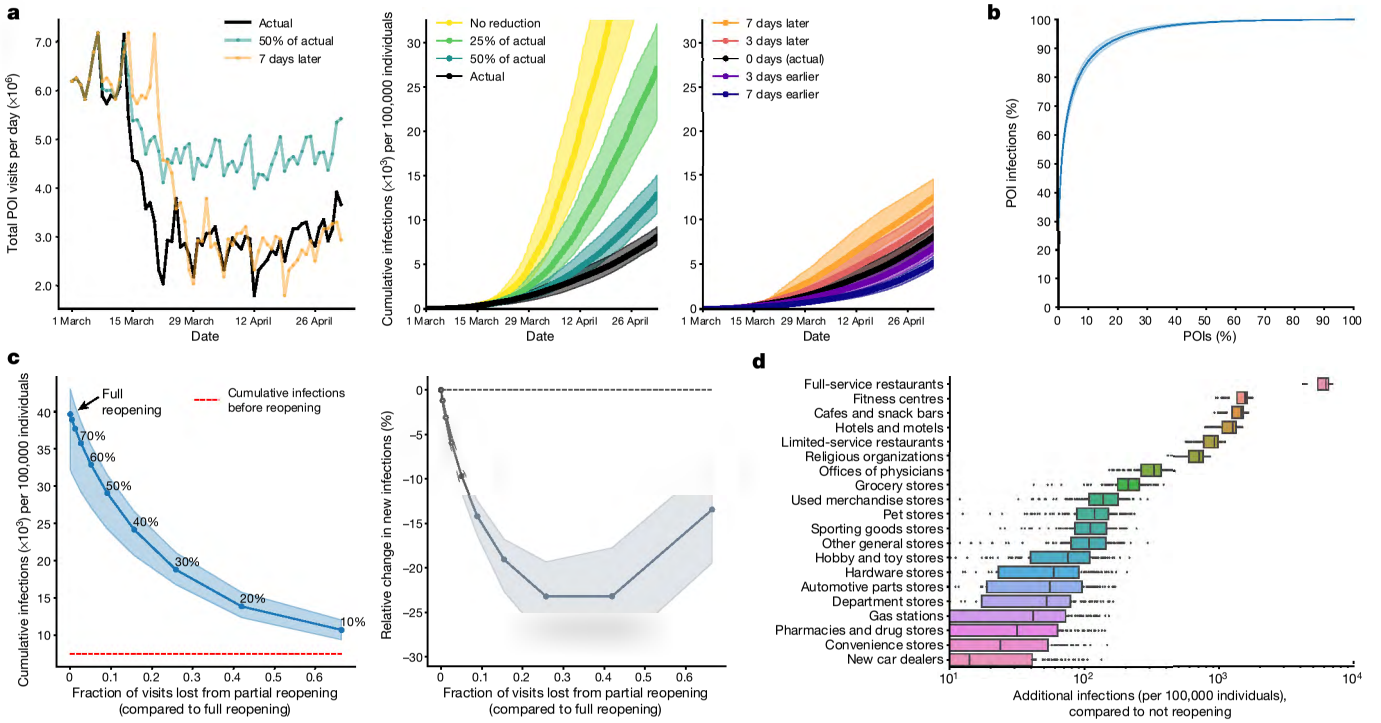


Fig. 2 | Assessing mobility reduction and reopening. The Chicago metro area is used as an example; results for all metro areas are included in Extended Data Figs. 3, 4, Supplementary Figs. 10, 15–24 and Supplementary Tables 4, 5, as indicated. **a**, Counterfactual simulations (left) of past reductions in mobility illustrate that the magnitude of the reduction (middle) was at least as important as its timing (right) (Supplementary Tables 4, 5). **b**, The model predicts that most infections at POIs occur at a small fraction of superspreader POIs (Supplementary Fig. 10). **c**, Left, the cumulative number of predicted infections after one month of reopening is plotted against the fraction of visits lost by partial instead of full reopening (Extended Data Fig. 3); the annotations within the plot show the fraction of maximum occupancy that is used as the cap and the horizontal red line indicates the cumulative number of predicted infections at the point of reopening (on 1 May 2020). Compared to full reopening, capping at 20% of the maximum occupancy in Chicago reduces the number of new infections by more than 80%, while only losing 42% of overall visits. Right, compared to uniformly reducing visits, the reduced maximum occupancy strategy always results in a smaller predicted increase in infections for the same number of visits (Extended Data Fig. 4). The horizontal grey line at 0% indicates when the two strategies result in an equal number of infections, and we observe that the curve falls well below this baseline. The *y* axis plots the relative difference between the predicted number of new infections under the reduced occupancy strategy compared to a uniform reduction. **d**, Reopening full-service restaurants has the largest predicted impact on infections, due to the large number of restaurants as well as their high visit densities and long dwell times (Supplementary Figs. 15–24). Colours are used to distinguish the different POI categories, but do not have any additional meaning. All results in this figure are aggregated across 4 parameter sets and 30 stochastic realizations ($n = 120$). Shaded regions in **a**–**c** denote the 2.5th to 97.5th percentiles; boxes in **d** denote the interquartile range and data points outside this range are shown as individual dots.

each policy, and running our SEIR model forward with this hypothetical network. Using this approach, we assess a wide range of mobility reduction and reopening strategies.

The magnitude of mobility reduction is at least as important as its timing

Mobility in the USA dropped sharply in March 2020: for example, overall POI visits in the Chicago metro area fell by 54.7% between the first week of March and the first week of April 2020. We constructed counterfactual mobility networks by scaling the magnitude of mobility reduction down and by shifting the timeline earlier and later, and applied our model to the counterfactual networks to simulate the resulting infection trajectories. Across metro areas, we found that the magnitude of mobility reduction was at least as important as its timing (Fig. 2a and Supplementary Tables 4, 5): for example, if the mobility reduction in the Chicago metro area had been only a quarter of the size, the predicted number of infections would have increased by 3.3× (95% confidence interval, 2.8–3.8×), compared with a 1.5× (95% confidence interval, 1.4–1.6×) increase had people begun reducing their mobility one full week later. Furthermore, if no mobility reduction had occurred at all, the predicted number of infections in the Chicago metro area would have increased by 6.2× (95% confidence interval, 5.2–7.1×). Our results are in accordance with previous findings that mobility reductions can markedly reduce infections[18,19,33,34].

A minority of POIs account for the majority of the predicted infections

We next investigated whether it matters how we reduce mobility—that is, to which POIs. We computed the number of infections that occurred at each POI in our simulations from 1 March to 2 May 2020, and found that the majority of the predicted infections occurred at a small fraction of superspreader POIs; for example, in the Chicago metro area, 10% of POIs accounted for 85% (95% confidence interval, 83–87%) of the predicted infections at the POIs (Fig. 2b and Supplementary Fig. 10). Certain categories of POIs also contributed far more to infections (for example, full-service restaurants and hotels), although our model predicted time-dependent variation in how much each category contributed (Extended Data Fig. 2). For example, restaurants and fitness centres contributed less to the predicted number of infections over time, probably because of lockdown orders to close these POIs, whereas grocery stores remained steady or even grew in their contribution, which is in agreement with their status as essential businesses.

Reopening with a reduced maximum occupancy

If a minority of POIs produce the majority of infections, then reopening strategies that specifically target high-risk POIs should be especially effective. To test one such strategy, we simulated reopening on 1 May, and modelled the effects of reducing the maximum occupancy in which

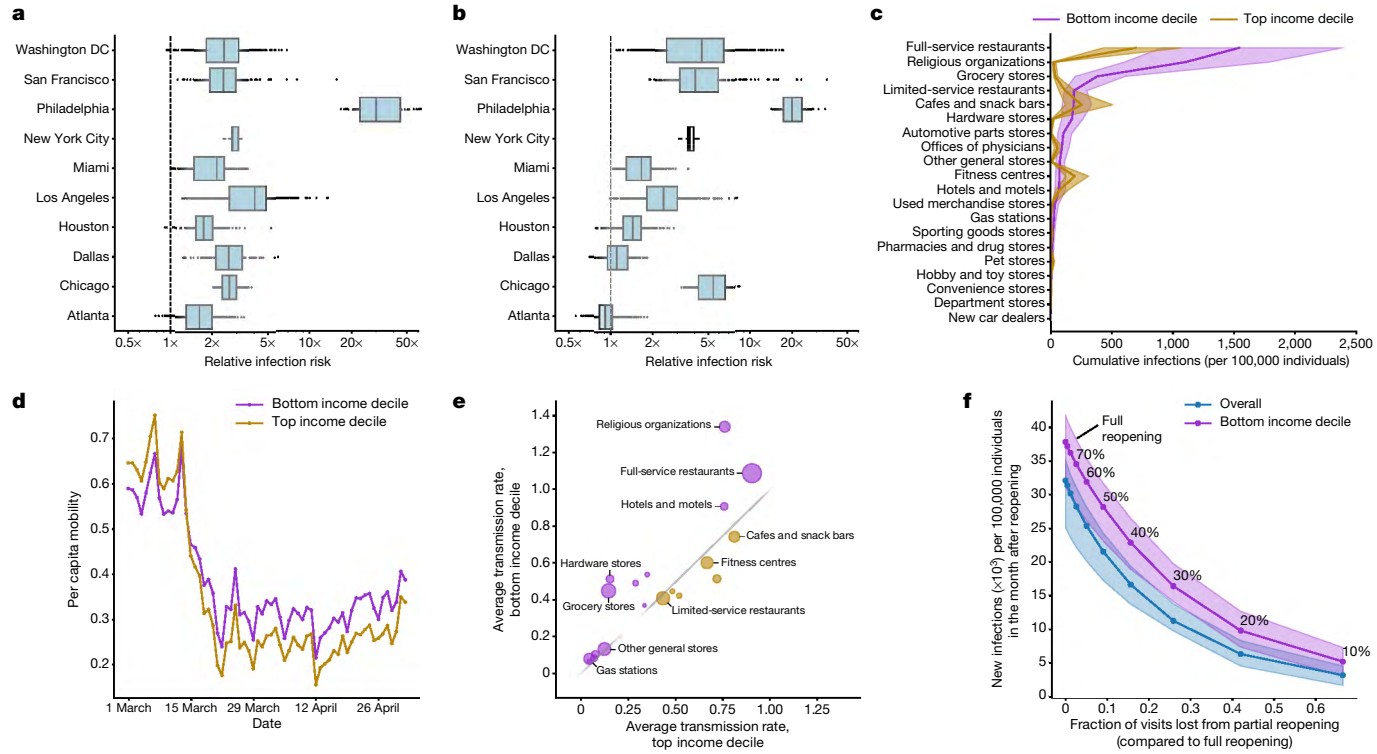


Fig. 3 | Mobility patterns give rise to infection disparities. a, In every metro area, our model predicts that people in lower-income CBGs are likelier to be infected. **b**, People in non-white CBGs area are also likelier to be infected, although results are more variable across metro areas. For **c**–**f**, the Chicago metro area is used as an example, but references to results for all metro areas are provided for each panel. **c**, The overall predicted disparity is driven by a few POI categories such as full-service restaurants (Supplementary Fig. 2). **d**, One reason for the predicted disparities is that higher-income CBGs were able to reduce their mobility levels below those of lower-income CBGs (Extended Data Fig. 6). **e**, Within each POI category, people from lower-income CBGs tend to visit POIs that have higher predicted transmission rates (Extended Data Table 3). The size of each dot represents the average number of visits per capita made to the category. The top 10 out of 20 categories with the most visits are labelled, covering 0.48–2.88 visits per capita (hardware stores–full-service restaurants). **f**, Reopening (at different levels of reduced maximum occupancy) leads to more predicted infections in lower-income CBGs than in the overall population (Extended Data Fig. 3). In **c**–**f**, purple denotes lower-income CBGs, yellow denotes higher-income CBGs and blue represents the overall population. Aside from **d** and **e**, which were directly extracted from mobility data, all results in this figure represent predictions aggregated over model realizations. Across metro areas, we sample 97 parameter sets, with 30 stochastic realizations each (n = 2,910); see Supplementary Table 6 for the number of sets per metro area. Shaded regions in **c** and **f** denote the 2.5th–97.5th percentiles; boxes in (**a**, **b**) denote the interquartile range; data points outside the range are shown as individual dots.

the numbers of hourly visits to each POI returned to their 'normal' levels from the first week of March but were capped if they exceeded a fraction of the maximum occupancy of that POI[35]. Full reopening without reducing the maximum occupancy produced a spike in the predicted number of infections: in the Chicago metro area, our models projected that an additional 32% (95% confidence interval, 25–35%) of the population would be infected by the end of May (Fig. 2c). However, reducing the maximum occupancy substantially reduced the risk without sharply reducing overall mobility: capping at 20% of the maximum occupancy in the Chicago metro area reduced the predicted number of new infections by more than 80% but only lost 42% of overall visits, and we observed similar trends across other metro areas (Extended Data Fig. 3). This result highlights the nonlinearity of the predicted number of infections as a function of the number of visits: one can achieve a disproportionately large reduction in infections with a small reduction in visits. Furthermore, in comparison to a different reopening strategy, in which the number of visits to each POI was uniformly reduced from their levels in early March, reducing the maximum occupancy always resulted in fewer predicted infections for the same number of total visits (Fig. 2c and Extended Data Fig. 4). This is because reducing the maximum occupancies takes advantage of the time-varying visit density within each POI, disproportionately reducing visits to the POI during the high-density periods with the highest risk, but leaving visit counts unchanged during periods with lower risks. These results support previous findings that precise interventions, such as reducing the maximum occupancy, may be more effective than less targeted measures, while incurring substantially lower economic costs[36].

Relative risk of reopening different categories of POIs

Because we found that certain POI categories contributed far more to predicted infections in March (Extended Data Fig. 2), we also expected that reopening some POI categories would be riskier than reopening others. To assess this, we simulated reopening each category in turn on 1 May 2020 (by returning its mobility patterns to early March levels, as above), while keeping all other POIs at their reduced mobility levels from the end of April. We found large variation in predicted reopening risks: on average across metro areas, full-service restaurants, gyms, hotels, cafes, religious organizations and limited-service restaurants produced the largest predicted increases in infections when reopened (Extended Data Fig. 5d). Reopening full-service restaurants was associated with a particularly high risk: in the Chicago metro area, we predicted an additional 595,805 (95% confidence interval, 433,735–685,959) infections by the end of May, more than triple that of the POI category with the next highest risk (Fig. 2d). These risks are summed over all POIs in the category, but the relative risks after normalizing by the number of POIs were broadly similar (Extended Data Fig. 5c). These categories were predicted to be have a higher risk because, in the mobility data, their POIs tended to have higher visit densities and/or visitors stayed there longer (Supplementary Figs. 15–24).

Demographic disparities in infections

We characterize the differential spread of SARS-CoV-2 along demographic lines by using US census data to annotate each CBG with its racial composition and median income, then tracking predicted infection rates in CBGs with different demographic compositions: for example, within each metro area, comparing CBGs in the top and bottom deciles for income. We use this approach to study the mobility mechanisms behind disparities and to quantify how different reopening strategies affect disadvantaged groups.

Predicting disparities from mobility data

Despite having access to only mobility data and no demographic information, our models correctly predicted higher risks of infection among disadvantaged racial and socioeconomic groups[2–8]. Across all metro areas, individuals from CBGs in the bottom decile for income had a substantially higher likelihood of being infected by the end of the simulation, even though all individuals began with equal likelihoods of infection (Fig. 3a). This predicted disparity was driven primarily by a few POI categories (for example, full-service restaurants); far greater proportions of individuals from lower-income CBGs than higher-income CBGs became infected in these POIs (Fig. 3c and Supplementary Fig. 2). We similarly found that CBGs with fewer white residents had higher predicted risks of infection, although results were more variable across metro areas (Fig. 3b). In the Supplementary Discussion, we confirm that the magnitude of the disparities that our model predicts is generally consistent with real-world disparities and further explore the large predicted disparities in Philadelphia, that stem from substantial differences in the POIs that are frequented by higher- versus lower-income CBGs. In the analysis below, we discuss two mechanisms that lead higher predicted infection rates among lower-income CBGs, and we show in Extended Data Fig. 6 and Extended Data Table 4 that similar results hold for racial disparities as well.

Lower-income CBGs saw smaller reductions in mobility

A first mechanism producing disparities was that, across all metro areas, lower-income CBGs did not reduce their mobility as sharply in the first few weeks of March 2020, and these groups showed higher mobility than higher-income CBGs for most of March–May (Fig. 3d and Extended Data Fig. 6). For example, in April, individuals from lower-income CBGs in the Chicago metro area had 27% more POI visits per capita than those from higher-income CBGs. Category-level differences in visit patterns partially explained the infection disparities within each category: for example, individuals from lower-income CBGs made substantially more visits per capita to grocery stores than did those from higher-income CBGs (Supplementary Fig. 3) and consequently experienced more predicted infections for that category (Supplementary Fig. 2).

POIs visited by lower-income CBGs have higher transmission rates

Differences in visits per capita do not fully explain the infection disparities: for example, cafes and snack bars were visited more frequently by higher-income CBGs in every metro area (Supplementary Fig. 3), but our model predicted that a larger proportion of individuals from lower-income CBGs were infected at cafes and snack bars in the majority of metro areas (Supplementary Fig. 2). We found that even within a POI category, the predicted transmission rates at POIs frequented by individuals fom lower-income CBGs tended to be higher than the corresponding rates for those from higher-income CBGs (Fig. 3e and Extended Data Table 3), because POIs frequented by individuals from lower-income CBGs tended to be smaller and more crowded in the mobility data. As a case study, we examined grocery stores in further detail. In eight of the ten metro areas, visitors from lower-income CBGs encountered higher predicted transmission rates at grocery stores than visitors from higher-income CBGs (median transmission rate ratio of 2.19) (Extended Data Table 3). We investigated why one visit to the grocery store was predicted to be twice as dangerous for an individual from a lower-income CBG: the mobility data showed that the average grocery store visited by individuals from lower-income CBGs had 59% more hourly visitors per square foot, and their visitors stayed 17% longer on average (medians across metro areas). These findings highlight how fine-grained differences in mobility patterns—how often people go out and which POIs that they go to—can ultimately contribute to marked disparities in predicted infection outcomes.

Reopening plans must account for disparate effects

Because disadvantaged groups suffer a larger burden of infection, it is critical to not only consider the overall impact of reopening plans but also their disparate effects on disadvantaged groups specifically. For example, our model predicted that full reopening in the Chicago metro area would result in an additional 39% (95% confidence interval, 31–42%) of the population of CBGs in the bottom income decile being infected within a month, compared to 32% (95% confidence interval, 25–35%) of the overall population (Fig. 3f; results for all metro areas are shown in Extended Data Fig. 3). Similarly, Supplementary Fig. 4 illustrates that reopening individual POI categories tends to have a larger predicted effect on lower-income CBGs. More stringent reopening plans produce smaller absolute disparities in predicted infections—for example, we predict that reopening at 20% of the maximum occupancy in Chicago would result in additional infections for 6% (95% confidence interval, 4–8%) of the overall population and 10% (95% confidence interval, 7–13%) of the population in CBGs in the bottom income decile (Fig. 3f)—although the relative disparity remains.

Discussion

The mobility dataset that we use has limitations: it does not cover all populations, does not contain all POIs and cannot capture sub-CBG heterogeneity. Our model itself is also parsimonious, and does not include all real-world features that are relevant to disease transmission. We discuss these limitations in more detail in the Supplementary Discussion. However, the predictive accuracy of our model suggests that it broadly captures the relationship between mobility and transmission, and we thus expect our broad conclusions—for example, that people from lower-income CBGs have higher infection rates in part because they tend to visit denser POIs and because they have not reduced mobility by as much (probably because they cannot work from home as easily[4])—to hold robustly. Our fine-grained network modelling approach naturally extends to other mobility datasets and models that capture more aspects of real-world transmission, and these represent interesting directions for future work.

Our results can guide policy-makers that seek to assess competing approaches to reopening. Despite growing concern about racial and socioeconomic disparities in infections and deaths, it has been difficult for policy-makers to act on those concerns; they are currently operating without much evidence on the disparate effects of reopening policies, prompting calls for research that both identifies the causes of observed disparities and suggests policy approaches to mitigate them[5,8,37,38]. Our fine-grained mobility modelling addresses both these needs. Our results suggest that infection disparities are not the unavoidable consequence of factors that are difficult to address in the short term, such as differences in preexisting conditions; on the contrary, short-term policy decisions can substantially affect infection outcomes by altering the overall amount of mobility allowed and the types of POIs reopened. Considering the disparate effects of reopening plans may lead policy-makers to adopt policies that can drive down infection densities in disadvantaged neighbourhoods by supporting, for example, more stringent caps on POI occupancies, emergency food distribution centres to reduce densities in high-risk stores, free and widely available testing in neighbourhoods predicted to be high risk

(especially given known disparities in access to tests[2]), improved paid leave policy or income support that enables essential workers to curtail mobility when sick, and improved workplace infection prevention for essential workers, such as high-quality personal protective equipment, good ventilation and physical distancing when possible. As reopening policies continue to be debated, it is critical to build tools that can assess the effectiveness and equity of different approaches. We hope that our model, by capturing heterogeneity across POIs, demographic groups and cities, helps to address this need.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2923-3.

1. Buckee, C. O. et al. Aggregated mobility data could help fight COVID-19. *Science* **368**, 145–146 (2020).
2. Wilson, C. These graphs show how COVID-19 is ravaging New York City's low-income neighborhoods. *Time* (15 April 2020).
3. Garg, S. et al. Hospitalization rates and characteristics of patients hospitalized with laboratory-confirmed coronavirus disease 2019 — COVID-NET, 14 states, March 1—30, 2020. *MMWR Morb. Mortal. Wkly Rep.* **69**, 458–464 (2020).
4. Reeves, R. V. & Rothwell, J. *Class and COVID: How the Less Affluent face Double Risks*. https://www.brookings.edu/blog/up-front/2020/03/27/class-and-covid-how-the-less-affluent-face-double-risks/ (The Brookings Institution, 2020).
5. Pareek, M. et al. Ethnicity and COVID-19: an urgent public health research priority. *Lancet* **395**, 1421–1422 (2020).
6. Dorn, A. V., Cooney, R. E. & Sabin, M. L. COVID-19 exacerbating inequalities in the US. *Lancet* **395**, 1243–1244 (2020).
7. Yancy, C. W. COVID-19 and African Americans. *J. Am. Med. Assoc.* **323**, 1891–1892 (2020).
8. Chowkwanyun, M. & Reed, A. L. Jr. Racial Health Disparities and Covid-19 — caution and context. *N. Engl. J. Med.* **383**, 201–203 (2020).
9. Flaxman, S. et al. Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe. *Nature* **584**, 257–261 (2020).
10. Rojas, R. & Delkic, M. As states reopen, governors balance existing risks with new ones. *The New York Times* (17 May 2020).
11. Endo, A., Abbott, S., Kucharski, A. J. & Funk, S. Estimating the overdispersion in COVID-19 transmission using outbreak sizes outside China. *Wellcome Open Res.* **5**, 67 (2020).
12. Adam, D. C. et al. Clustering and superspreading potential of SARS-CoV-2 infections in Hong Kong. *Nat. Med.* https://doi.org/10.1038/s41591-020-1092-0 (2020).
13. Park, S. Y. et al. Coronavirus disease outbreak in call center, South Korea. *Emerg. Infect. Dis.* **26**, 1666–1670 (2020).
14. Bi, Q. et al. Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study. *Lancet Infect. Dis.* **20**, 911–919 (2020).
15. Chinazzi, M. et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).
16. Jia, J. S. et al. Population flow drives spatio-temporal distribution of COVID-19 in China. *Nature* **582**, 389–394 (2020).
17. Pei, S., Kandula, S. & Shaman, J. Differential effects of intervention timing on COVID-19 spread in the United States. Preprint at https://doi.org/10.1101/2020.05.15.20103655 (2020).
18. Lai, S. et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature* **585**, 410–413 (2020).
19. Badr, H. S. et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. *Lancet Infect. Dis.* **20**, 1247–1254 (2020).
20. Li, R. et al. Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2). *Science* **368**, 489–493 (2020).
21. Pei, S. & Shaman, J. Initial simulation of SARS-CoV2 spread and intervention effects in the continental US. Preprint at https://doi.org/10.1101/2020.03.21.20040303 (2020).
22. Aleta, A. et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. *Nat. Hum. Behav.* **4**, 964–971 (2020).
23. Duque, D. et al. Timing social distancing to avert unmanageable COVID-19 hospital surges. *Proc. Natl Acad. Sci. USA* **177**, 19873–19878 (2020).
24. Block, P. et al. Social network-based distancing strategies to flatten the COVID-19 curve in a post-lockdown world. *Nat. Hum. Behav.* **4**, 588–596 (2020).
25. Karin, O. et al. Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity. Preprint at https://doi.org/10.1101/2020.04.04.20053579 (2020).
26. Gao, S. et al. Mapping county-level mobility pattern changes in the United States in response to COVID-19. *SIGSPATIAL Special* **12**, 16–26 (2020).
27. Klein, B. et al. *Assessing Changes in Commuting and Individual Mobility in Major Metropolitan Areas in the United States during the COVID-19 Outbreak*. https://www.networkscienceinstitute.org/publications/assessing-changes-in-commuting-and-individual-mobility-in-major-metropolitan-areas-in-the-united-states-during-the-covid-19-outbreak (2020).
28. Benzell, S. G., Collis, A. & Nicolaides, C. Rationing social contact during the COVID-19 pandemic: transmission risk and social benefits of US locations. *Proc. Natl Acad. Sci. USA* **117**, 14642–14644 (2020).
29. Baicker, K. et al. Is it safer to visit a coffee shop or a gym? *The New York Times* (6 May 2020).
30. Hsiang, S. et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. *Nature* **584**, 262–267 (2020).
31. Deming, W. E. & Stephan, F. F. On a least squares adjustment of a sampled frequency table when the expected marginal totals are known. *Ann. Math. Stat.* **11**, 427–444 (1940).
32. The New York Times. *Coronavirus (COVID-19) Data in the United States*. https://github.com/nytimes/covid-19-data (2020).
33. Tian, H. et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. *Science* **368**, 638–642 (2020).
34. Watts, D. J., Muhamad, R., Medina, D. C. & Dodds, P. S. Multiscale, resurgent epidemics in a hierarchical metapopulation model. *Proc. Natl Acad. Sci. USA* **102**, 11157–11162 (2005).
35. California Department of Public Health. *COVID-19 Industry Guidance: Retail*. https://covid19.ca.gov/pdf/guidance-retail.pdf (2020).
36. Birge, J., Candogan, O. & Feng, Y. Controlling epidemic spread: reducing economic losses with targeted closures. *BFI Working Paper No. 2020-57* (8 May 2020).
37. Webb Hooper, M., Nápoles, A. M. & Pérez-Stable, E. J. COVID-19 and racial/ethnic disparities. *J. Am. Med. Assoc.* **323**, 2466–2467 (2020).
38. Laurencin, C. T. & McClinton, A. The COVID-19 pandemic: a call to action to identify and address racial and ethnic disparities. *J. Racial Ethn. Health Disparities* **7**, 398–402 (2020).

Methods

The Methods is structured as follows. We describe the datasets that we used in the 'Datasets' section and the mobility network that we derived from these datasets in the 'Mobility network' section. In the 'Model dynamics' section, we discuss the SEIR model that we overlaid on the mobility network; in the 'Model calibration' section, we describe how we calibrated this model and quantified uncertainty in its predictions. Finally, in the 'Analysis details' section, we provide details on the experimental procedures used for our analyses of mobility reduction, reopening plans and demographic disparities.

Datasets

SafeGraph. We use data provided by SafeGraph, a company that aggregates anonymized location data from numerous mobile applications. SafeGraph data captures the movement of people between CBGs, which are geographical units that typically contain a population of between 600 and 3,000 people, and POIs such as restaurants, grocery stores or religious establishments. Specifically, we use the following SafeGraph datasets.

First, we used the Places Patterns[39] and Weekly Patterns (v1)[40] datasets. These datasets contain, for each POI, hourly counts of the number of visitors, estimates of median visit duration in minutes (the 'dwell time') and aggregated weekly and monthly estimates of the home CBGs of visitors. We use visitor home CBG data from the Places Patterns dataset: for privacy reasons, SafeGraph excludes a home CBG from this dataset if fewer than five devices were recorded at the POI from that CBG over the course of the month. For each POI, SafeGraph also provides their North American industry classification system category, as well as estimates of its physical area in square feet. The area is computed using the footprint polygon SafeGraph that assigns to the POI[41,42]. We analyse Places Patterns data from 1 January 2019 to 29 February 2020 and Weekly Patterns data from 1 March 2020 to 2 May 2020.

Second, we used the Social Distancing Metrics dataset[43], which contains daily estimates of the proportion of people staying home in each CBG. We analyse Social Distancing Metrics data from 1 March 2020 to 2 May 2020.

We focus on 10 of the largest metro areas in the United States (Extended Data Table 1). We chose these metro areas by taking a random subset of the SafeGraph Patterns data and selecting the 10 metro areas with the most POIs in the data. The application of the methods described in this paper to the other metro areas in the original SafeGraph data should be straightforward. For each metro area, we include all POIs that meet all of the following requirements: (1) the POI is located in the metro area ; (2) SafeGraph has visit data for this POI for every hour that we model, from 00:00 on 1 March 2020 to 23:00 on 2 May 2020; (3) SafeGraph has recorded the home CBGs of visitors to this POI for at least one month from January 2019 to February 2020; (4) the POI is not a 'parent' POI. Parent POIs comprise a small fraction of POIs in the dataset that overlap and include the visits from their 'child' POIs: for example, many malls in the dataset are parent POIs, which include the visits from stores that are their child POIs. To avoid double-counting visits, we remove all parent POIs from the dataset. After applying these POI filters, we include all CBGs that have at least one recorded visit to at least ten of the remaining POIs; this means that CBGs from outside the metro area may be included if they visit this metro area frequently enough. Summary statistics of the post-processed data are shown in Extended Data Table 1. Overall, we analyse 56,945 CBGs from the 10 metro areas, and more than 310 million visits from these CBGs to 552,758 POIs.

SafeGraph data have been used to study consumer preferences[44] and political polarization[45]. More recently, it has been used as one of the primary sources of mobility data in the USA for tracking the effects of the COVID-19 pandemic[26,28,46–48]. In Supplementary Methods section 1, we show that aggregate trends in SafeGraph mobility data match the aggregate trends in Google mobility data in the USA[49], before and after the imposition of stay-at-home measures. Previous analyses of SafeGraph data have shown that it is geographically representative—for example, it does not systematically overrepresent individuals from CBGs in different counties or with different racial compositions, income levels or educational levels[50,51].

US census. Our data on the demographics of the CBGs comes from the American Community Survey (ACS) of the US Census Bureau[52]. We use the 5-year ACS data (2013–2017) to extract the median household income, the proportion of white residents and the proportion of Black residents of each CBG. For the total population of each CBG, we use the most-recent one-year estimates (2018); one-year estimates are noisier but we wanted to minimize systematic downward bias in our total population counts (due to population growth) by making them as recent as possible.

***The New York Times* dataset.** We calibrated our models using the COVID-19 dataset published by the *The New York Times*[32]. Their dataset consists of cumulative counts of cases and deaths in the USA over time, at the state and county level. For each metro area that we modelled, we sum over the county-level counts to produce overall counts for the entire metro area. We convert the cumulative case and death counts to daily counts for the purposes of model calibration, as described in the 'Model calibration' section.

Data ethics. The dataset from *The New York Times* consists of aggregated COVID-19-confirmed case and death counts collected by journalists from public news conferences and public data releases. For the mobility data, consent was obtained by the third-party sources that collected the data. SafeGraph aggregates data from mobile applications that obtain opt-in consent from their users to collect anonymous location data. Google's mobility data consists of aggregated, anonymized sets of data from users who have chosen to turn on the location history setting. Additionally, we obtained IRB exemption for SafeGraph data from the Northwestern University IRB office.

Mobility network

Definition. We consider a complete undirected bipartite graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with time-varying edges. The vertices $\mathcal{V}$ are partitioned into two disjoint sets $\mathcal{C} = \{c_1, \ldots, c_m\}$, representing m CBGs, and $\mathcal{P} = \{p_1, \ldots, p_n\}$, representing n POIs. From US census data, each CBG c_i is labelled with its population N_{c_i}, income distribution, and racial and age demographics. From SafeGraph data, each POI p_j is similarly labelled with its category (for example, restaurant, grocery store or religious organization), its physical size in square feet a_{p_j}, and the median dwell time d_{p_j} of visitors to p_j. The weight $w_{ij}^{(t)}$ on an edge (c_i, p_j) at time t represents our estimate of the number of individuals from CBG c_i visiting POI p_j at the tth hour of simulation. We record the number of edges (with non-zero weights) in each metro area and for all hours from 1 March 2020 to 2 May 2020 in Extended Data Table 1. Across all 10 metro areas, we study 5.4 billion edges between 56,945 CBGs and 552,758 POIs.

Overview of the network estimation. The central technical challenge in constructing this network is estimating the network weights $W^{(t)} = \{w_{ij}^{(t)}\}$ from SafeGraph data, as this visit matrix is not directly available from the data. Our general methodology for network estimation is as follows.

First, from SafeGraph data, we derived a time-independent estimate $\bar{W}$ of the visit matrix that captures the aggregate distribution of visits from CBGs to POIs from January 2019 to February 2020.

Second, because visit patterns differ substantially from hour to hour (for example, day versus night) and day to day (for example, before versus after lockdown), we used current SafeGraph data to capture these hourly variations and to estimate the CBG marginals $U^{(t)}$, that

is, the number of people in each CBG who are out visiting POIs at hour t, as well as the POI marginals $V^{(t)}$, that is, the total number of visitors present at each POI p_j at hour t.

Finally, we applied the iterative proportional fitting procedure (IPFP) to estimate an hourly visit matrix $W^{(t)}$ that is consistent with the hourly marginals $U^{(t)}$ and $V^{(t)}$ but otherwise 'as similar as possible' to the distribution of visits in the aggregate visit matrix $\overline{W}$, in terms of Kullback–Leibler divergence.

IPFP is a classic statistical method[31] for adjusting joint distributions to match prespecified marginal distributions, and it is also known in the literature as biproportional fitting, the RAS algorithm or raking[53]. In the social sciences, it has been widely used to infer the characteristics of local subpopulations (for example, within each CBG) from aggregate data[54–56]. IPFP estimates the joint distribution of visits from CBGs to POIs by alternating between scaling each row to match the hourly row (CBG) marginals $U^{(t)}$ and scaling each column to match the hourly column (POI) marginals $V^{(t)}$. Further details about the estimation procedure are provided in Supplementary Methods section 3.

Model dynamics

To model the spread of SARS-CoV-2, we overlay a metapopulation disease transmission model on the mobility network defined in the 'Mobility Network' section. The transmission model structure follows previous work[15,20] on epidemiological models of SARS-CoV-2 but incorporates a fine-grained mobility network into the calculations of the transmission rate. We construct separate mobility networks and models for each metropolitan statistical area.

We use a SEIR model with susceptible (S), exposed (E), infectious (I) and removed (R) compartments. Susceptible individuals have never been infected, but can acquire the virus through contact with infectious individuals, which may happen at POIs or in their home CBG. They then enter the exposed state, during which they have been infected but are not infectious yet. Individuals transition from exposed to infectious at a rate inversely proportional to the mean latency period. Finally, they transition into the removed state at a rate inversely proportional to the mean infectious period. The removed state represents individuals who can no longer be infected or infect others, for example, because they have recovered, self-isolated or died.

Each CBG c_i maintains its own SEIR instantiation, with $S_{c_i}^{(t)}$, $E_{c_i}^{(t)}$, $I_{c_i}^{(t)}$ and $R_{c_i}^{(t)}$ representing how many individuals in CBG c_i are in each disease state at hour t, and $N_{c_i} = S_{c_i}^{(t)} + E_{c_i}^{(t)} + I_{c_i}^{(t)} + R_{c_i}^{(t)}$. At each hour t, we sample the transitions between states as follows:

$$N_{S_{c_i}\to E_{c_i}}^{(t)} \sim \text{Pois}\left(\frac{S_{c_i}^{(t)}}{N_{c_i}}\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)}\right) + \text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right) \quad (1)$$

$$N_{E_{c_i}\to I_{c_i}}^{(t)} \sim \text{Binom}\left(E_{c_i}^{(t)}, 1/\delta_E\right) \quad (2)$$

$$N_{I_{c_i}\to R_{c_i}}^{(t)} \sim \text{Binom}\left(I_{c_i}^{(t)}, 1/\delta_I\right), \quad (3)$$

where $\lambda_{p_j}^{(t)}$ is the rate of infection at POI p_j at time t; $w_{ij}^{(t)}$, the ijth entry of the visit matrix from the mobility network (see 'Mobility network'), is the number of visitors from CBG c_i to POI p_j at time t; $\lambda_{c_i}^{(t)}$ is the base rate of infection that is independent of visiting POIs; δ_E is the mean latency period; and δ_I is the mean infectious period.

We then update each state to reflect these transitions. Let $\Delta S_{c_i}^{(t)} = S_{c_i}^{(t+1)} - S_{c_i}^{(t)}$ and likewise for $\Delta E_{c_i}^{(t)}$, $\Delta I_{c_i}^{(t)}$ and $\Delta R_{c_i}^{(t)}$. Then, we make the following updates:

$$\Delta S_{c_i}^{(t)} = -N_{S_{c_i}\to E_{c_i}}^{(t)} \quad (4)$$

$$\Delta E_{c_i}^{(t)} = N_{S_{c_i}\to E_{c_i}}^{(t)} - N_{E_{c_i}\to I_{c_i}}^{(t)} \quad (5)$$

$$\Delta I_{c_i}^{(t)} = N_{E_{c_i}\to I_{c_i}}^{(t)} - N_{I_{c_i}\to R_{c_i}}^{(t)} \quad (6)$$

$$\Delta R_{c_i}^{(t)} = N_{I_{c_i}\to R_{c_i}}^{(t)}. \quad (7)$$

The number of new exposures. We separate the number of new exposures $N_{S_{c_i}\to E_{c_i}}^{(t)}$ in CBG c_i at time t into two parts: cases from visiting POIs, which are sampled from $\text{Pois}((S_{c_i}^{(t)}/N_{c_i})\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)})$, and other cases not captured by visiting POIs, which are sampled from $\text{Binom}(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)})$.

First, we calculate the number of new exposures from visiting POIs. We assume that any susceptible visitor to POI p_j at time t has the same independent probability $\lambda_{p_j}^{(t)}$ of being infected and transitioning from the susceptible (S) to the exposed (E) state. As there are $w_{ij}^{(t)}$ visitors from CBG c_i to POI p_j at time t, and we assume that a $S_{c_i}^{(t)}/N_{c_i}$ fraction of them are susceptible, the number of new exposures among these visitors is distributed as $\text{binom}(w_{ij}^{(t)}S_{c_i}^{(t)}/N_{c_i}, \lambda_{p_j}^{(t)}) \approx \text{Pois}(\lambda_{p_j}^{(t)}w_{ij}^{(t)}S_{c_i}^{(t)}/N_{c_i})$. The number of new exposures among all outgoing visitors from CBG c_i is therefore distributed as the sum of the above expression over all POIs, $\text{Pois}((S_{c_i}^{(t)}/N_{c_i})\sum_{j=1}^{n}\lambda_{p_j}^{(t)}w_{ij}^{(t)})$.

We model the infection rate at POI p_j at time t, $\lambda_{p_j}^{(t)} = \beta_{p_j}^{(t)}(I_{p_j}^{(t)}/V_{p_j}^{(t)})$, as the product of its transmission rate $\beta_{p_j}^{(t)}$ and proportion of infectious individuals $I_{p_j}^{(t)}/V_{p_j}^{(t)}$, where $V_{p_j}^{(t)} = \sum_{i=1}^{m} w_{ij}^{(t)}$ is the total number of visitors to p_j at time t. We model the transmission rate at POI p_j at time t as

$$\beta_{p_j}^{(t)} = \psi d_{p_j}^2 \frac{V_{p_j}^{(t)}}{a_{p_j}}, \quad (8)$$

where a_{p_j} is the physical area of p_j, and ψ is a transmission constant (shared across all POIs) that we fit to data. The inverse scaling of transmission rate with area a_{p_j} is a standard simplifying assumption[57]. The dwell time fraction $d_{p_j} \in [0, 1]$ is what fraction of an hour an average visitor to p_j at any hour will spend there (Supplementary Methods section 3); it has a quadratic effect on the POI transmission rate $\beta_{p_j}^{(t)}$ because it reduces both the time that a susceptible visitor spends at p_j and the density of visitors at p_j. With this expression for the transmission rate $\beta_{p_j}^{(t)}$, we can calculate the infection rate at POI p_j at time t as

$$\lambda_{p_j}^{(t)} = \beta_{p_j}^{(t)}\frac{I_{p_j}^{(t)}}{V_{p_j}^{(t)}} = \psi d_{p_j}^2 \frac{I_{p_j}^{(t)}}{a_{p_j}}. \quad (9)$$

For sufficiently large values of ψ and a sufficiently large proportion of infected individuals, the expression above can sometimes exceed 1. To address this, we simply clip the infection rate to 1. However, this occurs very rarely for the parameter settings and simulation duration that we use.

Finally, to compute the number of infectious individuals at p_j at time t, $I_{p_j}^{(t)}$, we assume that the proportion of infectious individuals among the $w_{kj}^{(t)}$ visitors to p_j from a CBG c_k mirrors the overall density of infections $I_{c_k}^{(t)}/N_{c_k}$ in that CBG, although we note that the scaling factor ψ can account for differences in the ratio of infectious individuals who visit POIs. This gives

$$I_{p_j}^{(t)} = \sum_{k=1}^{m}\frac{I_{c_k}^{(t)}}{N_{c_k}}w_{kj}^{(t)}. \quad (10)$$

In addition to the new exposures from infections at POIs, we model a CBG-specific base rate of new exposures that is independent of POI visit activity. This captures other sources of infections, for example, household infections or infections at POIs that are absent from the SafeGraph data. We assume that at each hour, every susceptible

individual in CBG c_i has a base probability $\lambda_{c_i}^{(t)}$ of becoming infected and transitioning to the exposed state, where

$$\lambda_{c_i}^{(t)} = \beta_{\text{base}} \frac{I_{c_i}^{(t)}}{N_{c_i}} \tag{11}$$

is the product of the base transmission rate β_{base} and the proportion of infectious individuals in CBG c_i. β_{base} is a constant (shared across all CBGs) that we fit to data.

By adding all of the above together, the expression for the distribution of the overall number of new exposures in CBG c_i at time t becomes

$$\begin{aligned} N_{S_{c_i}\to E_{c_i}}^{(t)} &\sim \text{Pois}\left(\frac{S_{c_i}^{(t)}}{N_{c_i}} \sum_{j=1}^{n} \lambda_{p_j}^{(t)} w_{ij}^{(t)}\right) + \text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right) \\ &= \underbrace{\text{Pois}\left(\psi \frac{S_{c_i}^{(t)}}{N_{c_i}} \sum_{j=1}^{n} \frac{d_{p_j}^2}{a_{p_j}} \left(\sum_{k=1}^{m} \frac{I_{c_k}^{(t)}}{N_{c_k}} w_{kj}^{(t)}\right) w_{ij}^{(t)}\right)}_{\text{New infections from visiting POIs}} \\ &\quad + \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \beta_{\text{base}} \frac{I_{c_i}^{(t)}}{N_{c_i}}\right)}_{\text{Base rate of new CBG infections}}. \end{aligned} \tag{12}$$

The number of new infectious and removed cases. We model exposed individuals as becoming infectious at a rate that is inversely proportional to the mean latency period δ_E. At each time step t, we assume that each exposed individual has a constant, time-independent probability of becoming infectious, with

$$N_{E_{c_i}\to I_{c_i}}^{(t)} \sim \text{Binom}\left(E_{c_i}^{(t)}, 1/\delta_E\right). \tag{13}$$

Similarly, we model infectious individuals as transitioning to the removed state at a rate that is inversely proportional to the mean infectious period δ_I, with

$$N_{I_{c_i}\to R_{c_i}}^{(t)} \sim \text{Binom}\left(I_{c_i}^{(t)}, 1/\delta_I\right), \tag{14}$$

We estimate $\delta_E = 96$ h (refs. [20,58]) and $\delta_I = 84$ h (ref. [20]) based on previous studies.

Model initialization. In our experiments, $t = 0$ is the first hour of 1 March 2020. We approximate the infectious I and removed R compartments at $t = 0$ as initially empty, with all infected individuals in the exposed E compartment. We further assume that the same expected initial prevalence p_0 occurs in every CBG c_i. At $t = 0$, every individual in the metro area has the same independent probability p_0 of being exposed E instead of susceptible S. We thus initialize the model state by setting

$$S_{c_i}^{(0)} = N_{c_i} - E_{c_i}^{(0)} \tag{15}$$

$$E_{c_i}^{(0)} \sim \text{Binom}\left(N_{c_i}, p_0\right) \tag{16}$$

$$I_{c_i}^{(0)} = 0 \tag{17}$$

$$R_{c_i}^{(0)} = 0. \tag{18}$$

Aggregate mobility and no-mobility baseline models. To test whether the detailed mobility network is necessary, or whether our model is simply making use of aggregate mobility patterns, we tested an alternative SEIR model that uses the aggregate number of visits made to any POI in the metro area in each hour, but not the breakdown of visits between specific CBGs to specific POIs. Like our model, the aggregate mobility model also captures new cases from visiting POIs and a base rate of infection that is independent of POI visit activity; thus, the two models have the same three free parameters (ψ, scaling transmission rates at POIs; β_{base}, the base transmission rate ; and p_0, the initial fraction of infected individuals). However, instead of having POI-specific rates of infection, the aggregate mobility model captures only a single probability that a susceptible person from any CBG will become infected due to a visit to any POI at time t; we make this simplification because the aggregate mobility model no longer has access to the breakdown of visits between CBGs and POIs. This probability $\lambda_{\text{POI}}^{(t)}$ is defined as

$$\lambda_{\text{POI}}^{(t)} = \psi \underbrace{\frac{\sum_{i=1}^{m}\sum_{j=1}^{n} w_{ij}^{(t)}}{nm}}_{\text{Average mobility at time } t} \frac{I^{(t)}}{N}, \tag{19}$$

where m is the number of CBGs, n is the number of POIs, $I^{(t)}$ is the total number of infectious individuals at time t, and N is the total population size of the metro area. For the base rate of infections in CBGs, we assume the same process as in our network model: the probability $\lambda_{c_i}^{(t)}$ that a susceptible person in CBG c_i will become infected in their CBG at time t is equal to β_{base} times the current infectious fraction of c_i (equation (11)). Putting it together, the aggregate mobility model defines the number of new exposures in CBG c_i at time t as

$$N_{S_{c_i}\to E_{c_i}}^{(t)} \sim \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \lambda_{\text{POI}}^{(t)}\right)}_{\text{New infections from visiting POIs}} + \underbrace{\text{Binom}\left(S_{c_i}^{(t)}, \lambda_{c_i}^{(t)}\right)}_{\text{Base rateof new CBG infections}}. \tag{20}$$

All other dynamics remain the same between the aggregate mobility model and our network model, and we calibrated the models in the same way, which we describe in the 'Model calibration' section. We found that our network model substantially outperformed the aggregate mobility model in predictions of out-of-sample cases: on average across metro areas, the out-of-sample r.m.s.e. of our best-fit network model was only 58% that of the best-fit aggregate mobility model (Extended Data Fig. 1). This demonstrates that it is not only general mobility patterns, but specifically the mobility network that allows our model to accurately fit observed cases.

Next, to determine the extent to which mobility data could aid in modelling the case trajectory, we compared our model to a baseline SEIR model that does not use mobility data and simply assumes that all individuals within an metro area mix uniformly. In this no-mobility baseline, an individual's risk of being infected and transitioning to the exposed state at time t is

$$\lambda^{(t)} = \beta_{\text{base}} \frac{I^{(t)}}{N}, \tag{21}$$

where $I^{(t)}$ is the total number of infectious individuals at time t, and N is the total population size of the metro area. As above, the other model dynamics are identical, and for model calibration we performed a similar grid search over β_{base} and p_0. As expected, we found both the network and aggregate mobility models outperformed the no-mobility model on out-of-sample case predictions (Extended Data Fig. 1).

Model calibration and validation

Most of our model parameters can either be estimated from SafeGraph and US census data, or taken from previous studies (see Extended Data Table 2 for a summary). This leaves three model parameters that do not have direct analogues in the literature, and that we therefore need to calibrate with data: (1) the transmission constant in POIs, ψ (equation (9)); (2) the base transmission rate, β_{base} (equation (11)); and (3) the initial proportion of exposed individuals at time $t = 0$, p_0 (equation (16)).

In this section, we describe how we fitted these parameters to published numbers of confirmed cases, as reported by *The New York Times*. We fitted models for each metro area separately.

Selecting parameter ranges for ψ, β_{base} and p_0. We select parameter ranges for the transmission rate factors ψ and β_{base} by checking whether the model outputs match plausible ranges of the basic reproduction number R_0 before lockdown, as R_0 has been the study of substantial previous work on SARS-CoV-2[59]. Under our model, we can decompose $R_0 = R_{base} + R_{POI}$, where R_{POI} describes transmission due to POIs and R_{base} describes the remaining transmission (as in equation (12)). We first establish plausible ranges for R_{base} and R_{POI} before translating these into plausible ranges for β_{base} and ψ.

We assume that R_{base} ranges from 0.1 to 2. R_{base} models transmission that is not correlated with activity at POIs in the SafeGraph dataset, including within-household transmission and transmission at POI categories that are not well-captured in the SafeGraph dataset. We chose the lower limit of 0.1 because beyond that point, base transmission would only contribute minimally to overall R, whereas previous studies have suggested that within-household transmission is a substantial contributor to overall transmission[60–62]. Household transmission alone is not estimated to be sufficient to tip the overall R_0 above 1; for example, a single infected individual has been estimated to cause an average of 0.32 (0.22–0.42) secondary within-household infections[60]. However, because R_{base} may also capture transmission at POIs that are not captured in the SafeGraph dataset, to be conservative, we chose an upper limit of $R_{base} = 2$; as we describe below, the best-fit models for all 10 metro areas have $R_{base} < 2$, and 9 out of 10 have $R_{base} < 1$. We allow R_{POI} to range from 1 to 3, which corresponds to allowing $R_0 = R_{POI} + R_{base}$ to range from 1.1 to 5. This is a conservatively wide range, as a previous study[59] estimated a pre-lockdown R_0 of 2–3.

To determine the values of R_{base} and R_{POI} that a given pair of β_{base} and ψ imply, we seeded a fraction of index cases and then ran the model on looped mobility data from the first week of March to capture pre-lockdown conditions. We initialized the model by setting p_0, the initial proportion of exposed individuals at time $t = 0$, to $p_0 = 10^{-4}$, and then sampling in accordance with equation (16). Let N_0 be the number of initial exposed individuals sampled. We computed the number of individuals that these N_0 index cases went on to infect through base transmission, N_{base}, and POI transmission, N_{POI}, which gives

$$R_{POI} = \frac{N_{POI}}{N_0} \tag{22}$$

$$R_{base} = \frac{N_{base}}{N_0}. \tag{23}$$

We averaged these quantities over stochastic realizations for each metro area. Supplementary Figure 6 shows that, as expected, R_{base} is linear in β_{base} and R_{POI} is linear in ψ. R_{base} lies in the plausible range when β_{base} ranges from 0.0012 to 0.024, and R_{POI} lies in the plausible range (for at least one metro area) when ψ ranges from 515 to 4,886; we therefore consider these parameter ranges when fitting the model.

The extent to which SARS-CoV-2 infections had spread in the USA by the start of our simulation (1 March 2020) is currently unclear[63]. To account for this uncertainty, we allow p_0 to vary across a large range between 10^{-5} and 10^{-2}. As described in the next section, we verified that case count data for all metro areas can be fit using parameter settings for β_{base}, ψ and p_0 within this range.

Fitting to the number of confirmed cases. Using the parameter ranges described above, we grid-searched over ψ, β_{base} and p_0 to find the models that best fit the number of confirmed cases reported by *The New York Times*[32]. For each metro area, we tested 1,500 different combinations of ψ, β_{base} and p_0 in the parameter ranges specified above, with parameters linearly spaced for ψ and β_{base} and logarithmically spread for p_0.

In the 'Model dynamics' section, we directly model the number of infections but not the number of confirmed cases. To estimate the number of confirmed cases, we assume that an $r_c = 0.1$ proportion[20,58,64–66] of infections will be confirmed, and moreover that they will confirmed exactly $\delta_c = 168$ h (7 days)[20,66] after becoming infectious. From these assumptions, we can calculate the predicted number of newly confirmed cases across all CBGs in the metro area on day d,

$$N_{cases}^{(\text{day } d)} = r_c \sum_{i=1}^{m} \sum_{\tau=24(d-1)+1-\delta_c}^{24d-\delta_c} N_{E_{c_i} \to I_{c_i}}^{(\tau)}, \tag{24}$$

where m indicates the total number of CBGs in the metro area and for convenience we define $N_{E_{c_i} \to I_{c_i}}^{(\tau)}$, the number of newly infectious people at hour τ, to be 0 when $\tau < 1$.

From the dataset of *The New York Times*, we have the reported number of new cases $\hat{N}_{cases}^{(\text{day } d)}$ for each day d, summed over each county in the metro area. We compare the reported number of cases and the number of cases that our model predicts by computing the r.m.s.e. between each of the $D = \lfloor T/24 \rfloor$ days of our simulations,

$$\text{r.m.s.e.} = \sqrt{\frac{1}{D} \sum_{d=1}^{D} \left(N_{cases}^{(\text{day } d)} - \hat{N}_{cases}^{(\text{day } d)} \right)^2}. \tag{25}$$

For each combination of model parameters and for each metro area, we quantify the model fit with the data from *The New York Times* by running 30 stochastic realizations and averaging their r.m.s.e. Note that we measure model fit based on the daily number of new reported cases (as opposed to the cumulative number of reported cases)[67].

Our simulation spans 1 March to 2 May 2020, and we use mobility data from that period. However, because we assume that cases will be confirmed $\delta_c = 7$ days after individuals become infectious, we predict the number of cases with a 7-day offset, from 8 March to 9 May 2020.

Parameter selection and uncertainty quantification. Throughout this paper, we report aggregate predictions from different parameter sets of ψ, β_{base} and p_0, and multiple stochastic realizations. For each metro area, we: (1) find the best-fit parameter set, that is, with the lowest average r.m.s.e. on daily incident cases over stochastic realizations; (2) select all parameter sets that achieve an r.m.s.e. (averaged over stochastic realizations) within 20% of the r.m.s.e. of the best-fit parameter set; and (3) pool together all predictions across those parameter sets and all of their stochastic realizations, and report their mean and 2.5–97.5th percentiles.

On average, each metro area has 9.7 parameter sets that achieve an r.m.s.e. within 20% of the best-fitting parameter set (Supplementary Table 6). For each parameter set, we have results for 30 stochastic realizations.

This procedure corresponds to rejection sampling in an approximate Bayesian computation framework[15], for which we assume an error model that is Gaussian with constant variance; we pick an acceptance threshold based on what the best-fit model achieves; and we use a uniform parameter grid instead of sampling from a uniform prior. It quantifies uncertainty from two sources. First, the multiple realizations capture stochastic variability between model runs with the same parameters. Second, simulating with all parameter sets that are within 20% of the r.m.s.e. of the best fit captures uncertainty in the model parameters ψ, β_{base} and p_0. This procedure is equivalent to assuming that the posterior probability over the true parameters is uniformly spread among all parameter sets within the 20% threshold.

Model validation on out-of-sample cases. We validate our models by showing that they predict the number of confirmed cases on out-of-sample data when we have access to corresponding mobility

data. For each metro area, we split the available dataset from the *The New York Times* into a training set (spanning from 8 March 2020 to 14 April 2020) and a test set (spanning from 15 April 2020 to 9 May 2020). We fit the model parameters ψ, β_{base} and p_0, as described in the 'Mobility network' section, but using only the training set. We then evaluate the predictive accuracy of the resulting model on the test set. When running our models on the test set, we still use mobility data from the test period. Thus, this is an evaluation of whether the models can accurately predict the number of cases, given mobility data, in a time period that was not used for model calibration. Extended Data Figure 1 shows that our network model fits the out-of-sample case data fairly well, and that our model substantially outperforms alternative models that use aggregated mobility data (without a network) or do not use mobility data at all (see 'Aggregate mobility and no-mobility baseline models'). Note that we only use this train/test split to evaluate out-of-sample model accuracy. All other results are generated using parameter sets that best fit the entire dataset, as described above.

Analysis details

In this section, we include additional details about the experiments that underlie the figures in the paper. We do not include explanations for figures that are completely described in the main text.

Counterfactuals of mobility reduction. Associated with Fig. 2a and Supplementary Tables 4, 5. To simulate what would have happened if we changed the magnitude or timing of mobility reduction, we modified the real mobility networks from 1 March to 2 May 2020, and then ran our models on the hypothetical data. In Fig. 2a, we report the total number of people per 100,000 of the population ever infected (that is, in the exposed, infectious and removed states) by the end of the simulation.

To simulate a smaller magnitude of mobility reduction, we interpolate between the mobility network from the first week of simulation (1–7 March 2020), which we use to represent typical mobility levels, and the actual observed mobility network for each week. Let $W^{(t)}$ represent the observed visit matrix at the tth hour of simulation, and let $f(t) = t \bmod 168$ map t to its corresponding hour in the first week of simulation, since there are 168 h in a week. To represent the scenario in which people had committed to $\alpha \in [0, 1]$ times the actual observed reduction in mobility, we construct a visit matrix $\widetilde{W}_{\alpha}^{(t)}$ that is an α-convex combination of $W^{(t)}$ and $W^{f(t)}$,

$$\widetilde{W}_{\alpha}^{(t)} = \alpha W^{(t)} + (1-\alpha) W^{f(t)}. \tag{26}$$

If α is 1, then $\widetilde{W}_{\alpha}^{(t)} = W^{(t)}$, and we use the actual observed mobility network for the simulation. On the other hand, if $\alpha = 0$, then $\widetilde{W}_{\alpha}^{(t)} = W^{f(t)}$, and we assume that people did not reduce their mobility levels at all by looping the visit matrix for the first week of March throughout the simulation. Any other $\alpha \in [0, 1]$ interpolates between these two extremes.

To simulate changing the timing of mobility reduction, we shift the mobility network by $d \in [-7, 7]$ days. Let T represent the last hour in our simulation (2 May 2020, 23:00), let $f(t) = t \bmod 168$ map t to its corresponding hour in the first week of simulation as above, and similarly let $g(t)$ map t to its corresponding hour in the last week of simulation (27 April–2 May 2020). We construct the time-shifted visit matrix $\widetilde{W}_{d}^{(t)}$

$$\widetilde{W}_{d}^{(t)} = \begin{cases} W^{(t-24d)} & \text{if } 0 \le t - 24d \le T, \\ W^{f(t-24d)} & \text{if } t - 24d < 0, \\ W^{g(t-24d)} & \text{otherwise.} \end{cases} \tag{27}$$

If d is positive, this corresponds to starting mobility reduction d days later; if we imagine time on a horizontal line, this shifts the time series to the right by $24d$ hours. However, doing so leaves the first $24d$ hours without visit data, so we fill it in by reusing visit data from the first week of simulation. Likewise, if d is negative, this corresponds to starting mobility reduction d days earlier, and we fill in the last $24d$ hours with visit data from the last week of simulation.

Distribution of predicted infections across POIs. Associated with Fig. 2b, Extended Data Fig. 2 and Supplementary Fig. 10. We run our models on the observed mobility data from 1 March–2 May 2020 and record the number of predicted infections that occur at each POI. Specifically, for each hour t, we compute the number of expected infections that occur at each POI p_j by taking the number of susceptible people who visit p_j in that hour multiplied by the POI infection rate $\lambda_{p_j}^{(t)}$ (equation (9)). In Fig. 2b and Supplementary Fig. 10, we sort the POIs by their total predicted number of infections (summed over hours) and plot the cumulative distribution of infections over this ordering of POIs. In Extended Data Fig. 2, we select the POI categories that contribute the most to predicted infections and plot the daily proportion of POI infections each category accounted for (summed over POIs within the category) over time.

Reducing mobility by capping maximum occupancy. Associated with Figs. 2c and Extended Data Fig. 3. We implemented two partial reopening strategies: one that uniformly reduced visits at POIs to a fraction of full activity, and the other that 'capped' the number of hourly visits at each POI to a fraction of the maximum occupancy of that the POI. For each reopening strategy, we started the simulation on 1 March 2020 and ran it until 30 May 2020, using the observed mobility network from 1 March to 30 April 2020, and then using a hypothetical post-reopening mobility network from 1 May to 30 May 2020, corresponding to the projected impact of that reopening strategy. Because we only have observed mobility data from 1 March to 2 May 2020, we impute the missing mobility data up to 30 May 2020 by looping mobility data from the first week of March, as in the above analysis on the effect of past reductions in mobility. Let T represent the last hour for which we have observed mobility data (2 May 2020, 23:00). To simplify the notation, we define

$$h(t) = \begin{cases} t & \text{if } t < T, \\ f(t) & \text{otherwise,} \end{cases} \tag{28}$$

where, as above, $f(t) = t \bmod 168$. This function leaves t unchanged if there is observed mobility data at time t, and otherwise maps t to the corresponding hour in the first week of our simulation.

To simulate a reopening strategy that uniformly reduced visits to an γ fraction of their original level, where $\gamma \in [0, 1]$, we constructed the visit matrix

$$\widetilde{W}_{\gamma}^{(t)} = \begin{cases} W^{h(t)} & \text{if } t < \tau, \\ \gamma W^{h(t)} & \text{otherwise,} \end{cases} \tag{29}$$

where τ represents the first hour of reopening (1 May 2020, 00:00). In other words, we use the actual observed mobility network up until hour τ, and then subsequently simulate an γ fraction of full mobility levels.

To simulate the reduced occupancy strategy, we first estimated the maximum occupancy M_{pj} of each POI p_j as the maximum number of visits that it ever had in one hour, across all of 1 March–2 May 2020. As in previous sections, let $w_{ij}^{(t)}$ represent the ijth entry in the observed visit matrix $W^{(t)}$, that is, the number of people from CBG c_i who visited p_j in hour t, and let $V_{p_j}^{(t)}$ represent the total number of visitors to p_j in that hour, that is, $\sum_i w_{ij}^{(t)}$. We simulated capping at a β fraction of maximum occupancy, where $\beta \in [0, 1]$, by constructing the visit matrix $\widetilde{W}_{\beta}^{(t)}$ for which the ijth entry is

$$\widetilde{w}_{ij\beta}^{(t)} = \begin{cases} w_{ij}^{h(t)} & \text{if } t < \tau \text{ or } V_{p_j}^{(t)} \le \beta M_{p_j}, \\ \dfrac{\beta M_{p_j}}{V_{p_j}^{(t)}} w_{ij}^{h(t)} & \text{otherwise.} \end{cases} \tag{30}$$

This corresponds to the following procedure: for each POI p_j and time t, we first check whether $t < \tau$ (reopening has not started) or whether $V_{p_j}^{(t)} \le \beta M_{p_j}$ (the total number of visits to p_j at time t is below the allowed maximum βM_{p_j}). If so, we leave $w_{ij}^{h(t)}$ unchanged. Otherwise, we compute the scaling factor $\frac{\beta M_{p_j}}{V_{p_j}^{(t)}}$ that would reduce the total visits to p_j at time t down to the allowed maximum βM_{p_j}, and then scale down all visits from each CBG c_i to p_j proportionately. For both reopening strategies, we calculate the predicted increase in cumulative incidence at the end of the reopening period (30 May 2020) compared to the start of the reopening period (1 May 2020).

Relative risk of reopening different categories of POIs. Associated with Fig. 2d, Extended Data Fig. 5 and Supplementary Figs. 11, 15–24. We study separately the reopening of the 20 POI categories with the most visits in SafeGraph data. In this analysis, we follow previous studies[28] and do not study four categories: 'child day-care services' and 'elementary and secondary schools' (because children under 13 are not well-tracked by SafeGraph); 'drinking places (alcoholic beverages)' (because SafeGraph seems to undercount these locations[28]) and 'nature parks and other similar institutions' (because boundaries and therefore areas are not well-defined by SafeGraph). We also exclude 'general medical and surgical hospitals' and 'other airport operations' (because hospitals and air travel both involve many additional risk factors that our model is not designed to capture). We do not filter out these POIs during model fitting (that is, we assume that people visit these POIs, and that transmissions occur there) because including them still increases the proportion of overall mobility that our dataset captures; we simply do not analyse these categories, because we wish to be conservative and only focus on categories for which we are most confident that we are capturing transmission faithfully.

This reopening analysis is similar to the previous experiments on reducing maximum occupancy versus uniform reopening (see 'Reducing mobility by capping maximum occupancy'). As above, we set the reopening time τ to 1 May 2020, 00:00. To simulate reopening a POI category, we take the set of POIs in that category, $\mathcal{V}$, and set their activity levels after reopening to that of the first week of March. For POIs not in the category $\mathcal{V}$, we keep their activity levels after reopening the same, that is, we simply repeat the activity levels of the last week of our data (27 April–2 May 2020): This gives us the visit matrix $\widetilde{W}^{(t)}$ with entries

$$\widetilde{w}_{ij}^{(t)} = \begin{cases} w_{ij}^{(t)} & \text{if } t < \tau, \\ w_{ij}^{f(t)} & \text{if } t \ge \tau, p_j \in \mathcal{V} \\ w_{ij}^{g(t)} & \text{if } t \ge \tau, p_j \notin \mathcal{V}. \end{cases} \quad (31)$$

As in the above reopening analysis, $f(t)$ maps t to the corresponding hour in the first week of March, and $g(t)$ maps t to the corresponding hour in the last week of our data. For each category, we calculate the predicted difference between (1) the cumulative fraction of people who have been infected by the end of the reopening period (30 May 2020) and (2) the cumulative fraction of people infected by 30 May 2020 had we not reopened the POI category (that is, if we simply repeated the activity levels of the last week of our data). This seeks to model the increase in cumulative incidence by the end of May from reopening the POI category. In Extended Data Fig. 5 and Supplementary Figs. 15–24, the bottom right panel shows the predicted increase for the category as a whole, and the bottom left panel shows the predicted increase per POI (that is, the total increase divided by the number of POIs in the category).

Per-capita mobility. Associated with Fig. 3d, Extended Data Fig. 6 and Supplementary Fig. 3. Each group of CBGs (for example, the bottom income decile) comprises a set $\mathcal{U}$ of CBGs that fit the corresponding criteria. In Fig. 3d and Extended Data Fig. 6, we show the daily per-capita mobilities of different pairs of groups (broken down by income and by race). To measure the per-capita mobility of a group on day d, we take the total number of visits made from those CBGs to any POI, $\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{P}} \sum_{t=24d}^{24d+23} w_{ij}^{(t)}$, and divide it by the total population of the CBGs in the group, $\sum_{c_i \in \mathcal{U}} N_{c_i}$. In Supplementary Fig. 3, we show the total number of visits made by each group to each POI category, accumulated over the entire data period (1 March–2 May 2020) and then divided by the total population of the group.

Average predicted transmission rate of a POI category. Associated with Fig. 3e and Extended Data Tables 3, 4. We compute the predicted average hourly transmission rate experienced by a group of CBGs $\mathcal{U}$ at a POI category $\mathcal{V}$ as

$$\overline{\beta}_{\mathcal{U}\mathcal{V}} = \frac{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t=1}^{T} w_{ij}^{(t)} \beta_{p_j}^{(t)}}{\sum_{c_i \in \mathcal{U}} \sum_{p_j \in \mathcal{V}} \sum_{t=1}^{T} w_{ij}^{(t)}}, \quad (32)$$

where, as above, $\beta_{p_j}^{(t)}$ is the transmission rate at POI p_j in hour t (equation (8)), $w_{ij}^{(t)}$ is the number of visitors from CBG c_i at POI p_j in hour t, and T is the last hour in our simulation. This represents the expected transmission rate encountered during a visit by someone from a CBG in group $\mathcal{U}$ to a POI in category $\mathcal{V}$.

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

Inferred hourly mobility networks as well as the estimated models are available at the project website (http://covid-mobility.stanford.edu). Raw census data (https://www.census.gov/programs-surveys/acs), case and death counts from *The New York Times* (https://github.com/nytimes/covid-19-data) and Google mobility data (https://www.google.com/covid19/mobility/) are also publicly available. Mobile phone mobility data are freely available to researchers, non-profit organizations and governments through the SafeGraph COVID-19 Data Consortium (https://www.safegraph.com/covid-19-data-consortium).

Code availability

Code is publicly available at the project website (http://covid-mobility.stanford.edu).

39. SafeGraph. *Places Schema*. https://docs.safegraph.com/docs/places-schema (2020).
40. SafeGraph. *Weekly Patterns*. https://docs.safegraph.com/docs/weekly-patterns (2020).
41. SafeGraph. *Using SafeGraph Polygons to Estimate Point-Of-Interest Square Footage*. https://www.safegraph.com/blog/using-safegraph-polygons-to-estimate-point-of-interest-square-footage (2019).
42. SafeGraph. *Guide to Points-of-Interest Data: POI Data FAQ*. https://www.safegraph.com/points-of-interest-poi-data-guide (2020).
43. SafeGraph. *Social Distancing Metrics*. https://docs.safegraph.com/docs/social-distancing-metrics (2020).
44. Athey, S. et al. Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. In *AEA Papers and Proceedings* Vol. 108, 64–67 (2018).
45. Chen, M. K. & Rohla, R. The effect of partisanship and political advertising on close family ties. *Science* **360**, 1020–1024 (2018).
46. Farboodi, M., Jarosch, G. & Shimer, R. Internal and external effects of social distancing in a pandemic. *NBER Working Paper 27059* https://doi.org/10.3386/w27059 (National Bureau Of Economic Research, 2020).
47. Killeen, B. D. et al. A county-level dataset for informing the United States' response to COVID-19. Preprint at https://arxiv.org/abs/2004.00756 (2020).
48. Allcott, H. et al. Polarization and public health: partisan differences in social distancing during the coronavirus pandemic. *NBER Working Paper 26946* https://doi.org/10.3386/w26946 (National Bureau Of Economic Research, 2020).

49. Google. *COVID-19 Community Mobility Reports*. https://google.com/covid19/mobility/ (2020).
50. Athey, S. et al. Experienced Segregation. *Working Paper 3785* https://gsb.stanford.edu/faculty-research/working-papers/experienced-segregation (2019).
51. Squire, R. F. *What about Bias in the SafeGraph Dataset?* https://safegraph.com/blog/what-about-bias-in-the-safegraph-dataset (2019).
52. US Census Bureau. *American Community Survey (ACS)* https://census.gov/programs-surveys/acs (published 30 January 2020).
53. Bishop, Y. M., Fienberg, S. E. & Holland, P. W. *Discrete Multivariate Analysis: Theory and Practice* (MIT Press, 1975).
54. Birkin, M. & Clarke, M. Synthesis—a synthetic spatial information system for urban and regional analysis: methods and examples. *Environ. Plann. A* **20**, 1645–1671 (1988).
55. Wong, D. W. The reliability of using the iterative proportional fitting procedure. *Prof. Geogr*. **44**, 340–348 (1992).
56. Simpson, L. & Tranmer, M. Combining sample and census data in small area estimates: iterative proportional fitting with standard software. *Prof. Geogr*. **57**, 222–234 (2005).
57. Hu, H., Nigmatulina, K. & Eckhoff, P. The scaling of contact rates with population density for the infectious disease models. *Math. Biosci*. **244**, 125–134 (2013).
58. Kucharski, A. J. et al. Early dynamics of transmission and control of COVID-19: a mathematical modelling study. *Lancet Infect. Dis*. **20**, 553–558 (2020).
59. Park, M., Cook, A. R., Lim, J. T., Sun, Y. & Dickens, B. L. A systematic review of COVID-19 epidemiology based on current evidence. *J. Clin. Med*. **9**, 967 (2020).
60. Curmei, M. et al. Estimating household transmission of SARS-CoV-2. Preprint at https://doi.org/10.1101/2020.05.23.20111559 (2020).
61. Li, W. et al. The characteristics of household transmission of COVID-19. *Clin. Infect. Dis*. **71**, 1943–1946 (2020).
62. Gudbjartsson, D. F. et al. Spread of SARS-CoV-2 in the Icelandic population. *N. Engl. J. Med*. **382**, 2302–2315 (2020).
63. Carey, B. & Glanz, J. Hidden outbreaks spread through U.S. cities far earlier than Americans knew, estimates say. *The New York Times* (23 April 2020).
64. Bommer, C. & Vollmer, S. *Average Detection Rate of SARS-CoV-2 Infections is Estimated Around Nine Percent*. https://www.uni-goettingen.de/en/606540.html (2020).
65. Javan, E., Fox, S. J. & Meyers, L. A. The unseen and pervasive threat of COVID-19 throughout the US. Preprint at https://doi.org/10.1101/2020.04.06.20053561 (2020).
66. Perkins, T. A. et al. Estimating unobserved SARS-CoV-2 infections in the United States. *Proc. Natl Acad. Sci. USA* **117**, 22597–22602 (2020).
67. King, A. A., Domenech de Cellès, M., Magpantay, F. M. & Rohani, P. Avoidable errors in the modelling of outbreaks of emerging pathogens, with special reference to Ebola. *Proc. R. Soc. Lond. B* **282**, 20150347 (2015).

Acknowledgements We thank Y.-Y. Ahn, R. Appel, C. Chen, J. Feng, N. Fishman, S. Fullerton, T. Hashimoto, M. Kraemer, P. Liang, M. Lipsitch, K. Loh, D. Ouyang, R. Rosenfeld, S. Sagawa, J. Steinhardt, R. Tibshirani, J. Ugander, D. Vrabac, seminar participants and Stanford's Computer Science and Civil Society for support and comments; and N. Singh, R. F. Squire, J. Williams-Holt, J. Wolf, R. Yang and others at SafeGraph for mobile phone mobility data and feedback. This research was supported by US National Science Foundation under OAC-1835598 (CINES), OAC-1934578 (HDR), CCF-1918940 (Expeditions), IIS-2030477 (RAPID), Stanford Data Science Initiative, Wu Tsai Neurosciences Institute and Chan Zuckerberg Biohub. S.C. was supported by an NSF Fellowship. E.P. was supported by a Hertz Fellowship. P.W.K. was supported by the Facebook Fellowship Program. J.L. is a Chan Zuckerberg Biohub investigator.

Author contributions S.C., E.P. and P.W.K. performed computational analysis. All authors jointly analysed the results and wrote the paper.

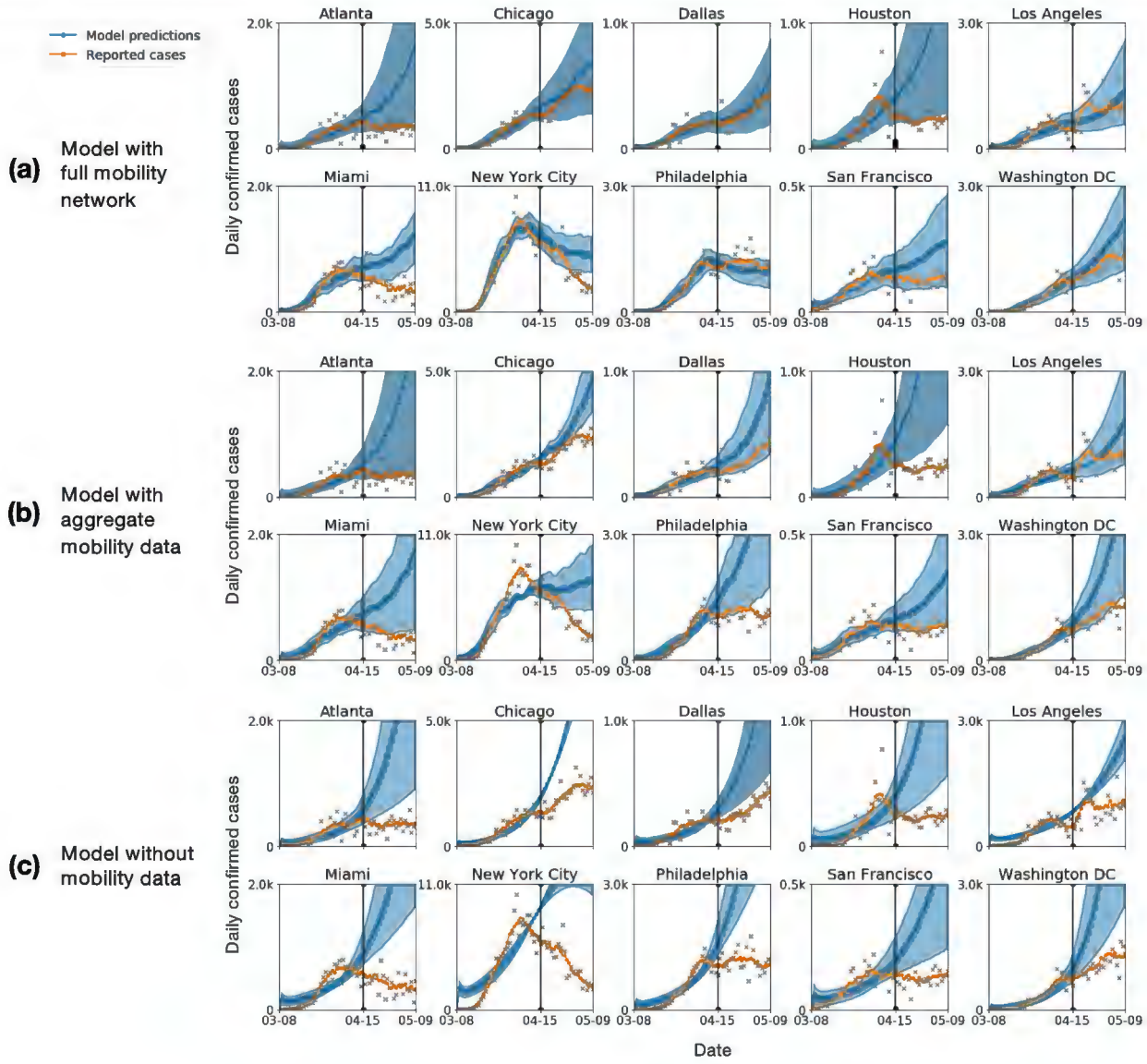
Competing interests The authors declare no competing interests.

Additional information

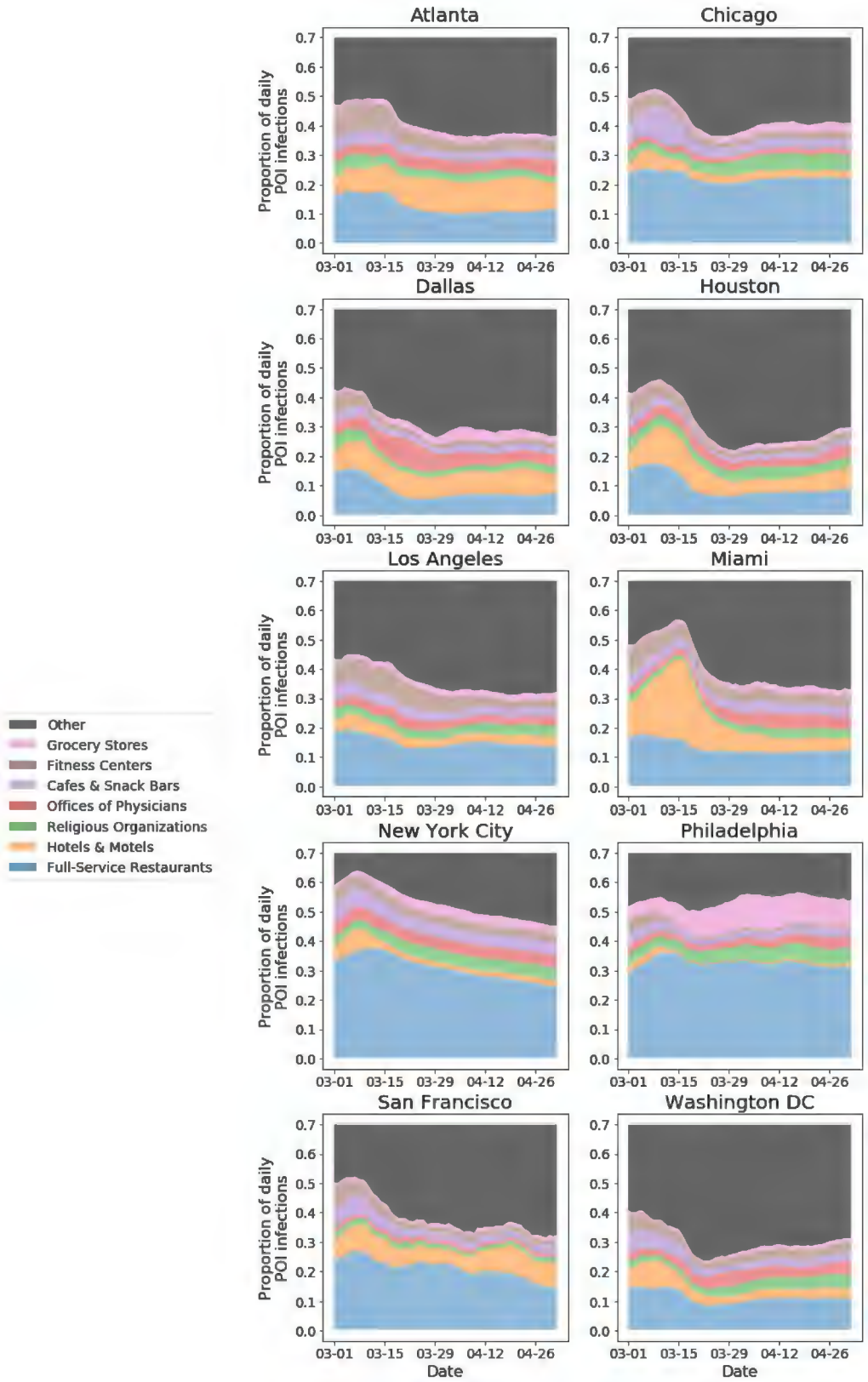
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2923-3.

Correspondence and requests for materials should be addressed to J.L.

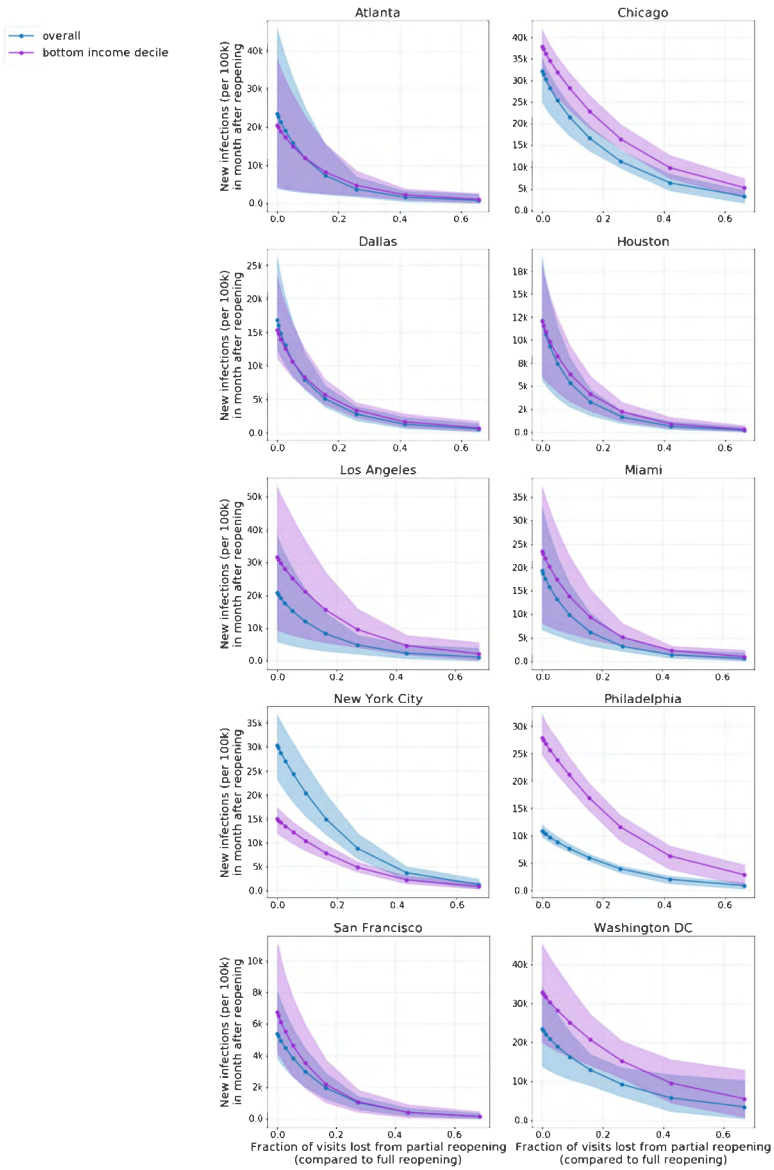
Peer review information *Nature* thanks Moritz Kraemer, Marc Lipsitch and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.



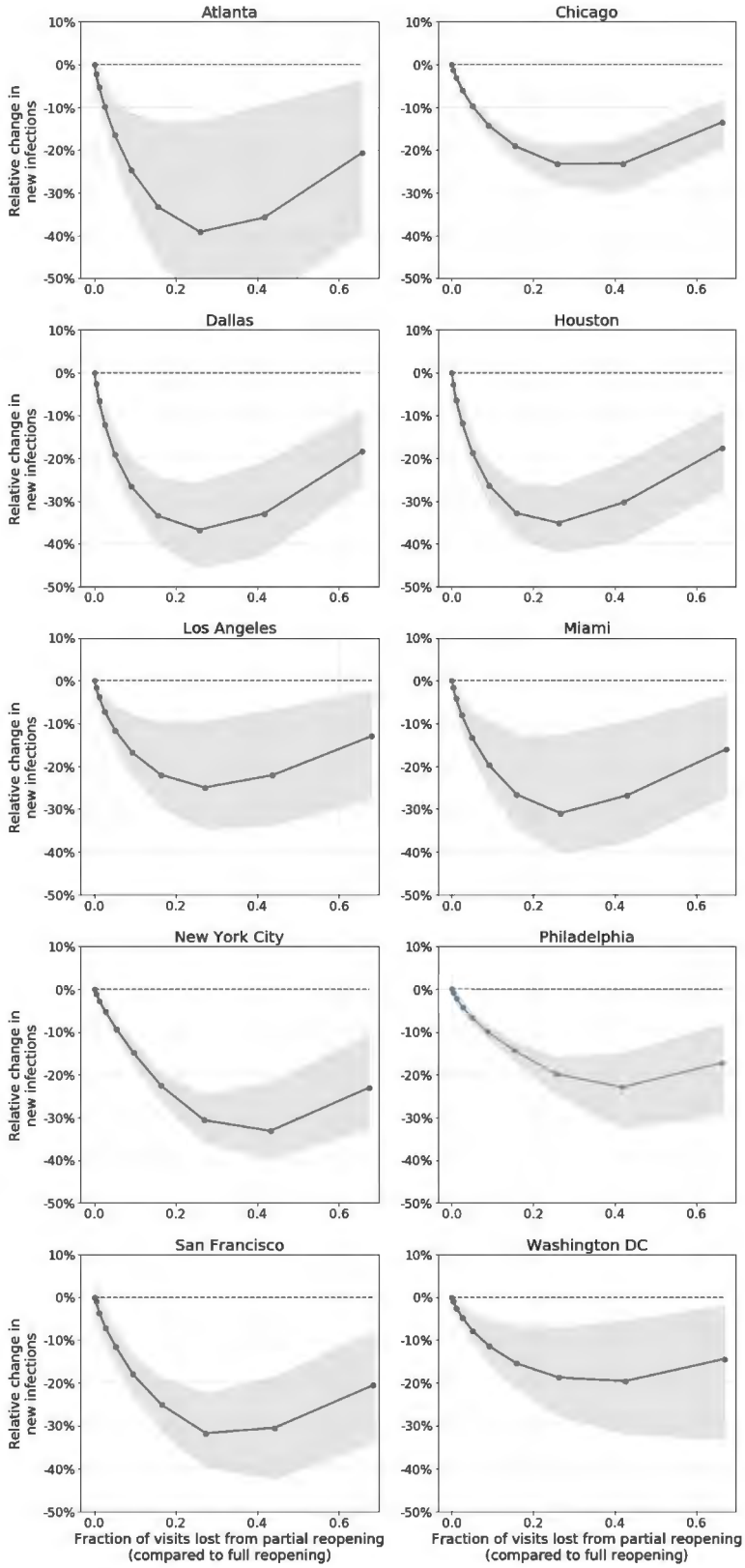
Extended Data Fig. 1 | Mobility-based epidemiological model and its predictions. a–**c**, Predicted (blue) and true (orange) daily case counts for our model (**a**), which uses hourly mobility networks, an SEIR model (**b**) that uses hourly aggregated mobility data and a baseline SEIR model (**c**) that does not use mobility data (see Methods, 'Aggregate mobility and no-mobility baseline models' for details). Incorporating mobility information improves out-of-sample fit and using a mobility network, instead of an aggregate measure of mobility, further improves fit: on average across metro areas, the out-of-sample error (r.m.s.e.) of our best-fit network model was only 58% that of the best-fit aggregate mobility model. All three models are calibrated on observed case counts before 15 April 2020 (vertical black line). The grey crosses represent the daily reported cases; as they tend to have great variability, we also show the smoothed weekly average (orange line). Shaded regions denote the 2.5th and 97.5th percentiles across sampled parameters and stochastic realizations.



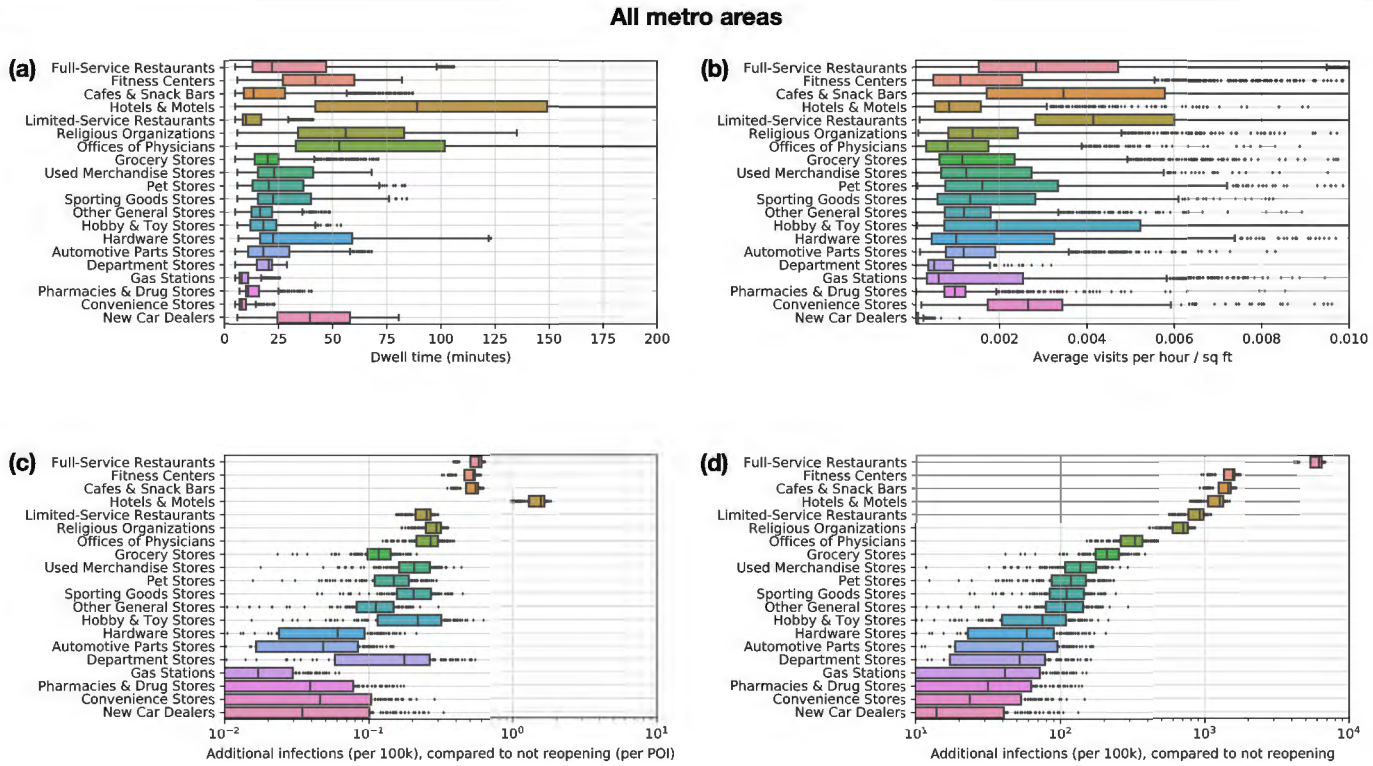
Extended Data Fig. 2 | Distribution of POI infections over time. We selected the POI categories that our models predicted to contribute the most to infections, and plotted the predicted proportion of POI infections that each category accounted for over time. Our model predicts time-dependent variation of where transmissions may have occurred. For example, full-service restaurants (blue) and fitness centres (brown) contributed less to predicted infections over time, probably due to lockdown orders closing these POIs, whereas grocery stores remained steady or even grew in their predicted contribution, probably because they remained open as essential businesses. Hotels and motels (yellow) also feature in these plots; most notably, the model predicts a peak in their contributed infections in Miami around mid-March, which aligns with college spring break, as Miami is a popular vacation spot for students. The proportions are stacked in these plots, and the *y* axes are truncated at 0.7 because every plot would only show 'other' from 0.7 to 1.0.



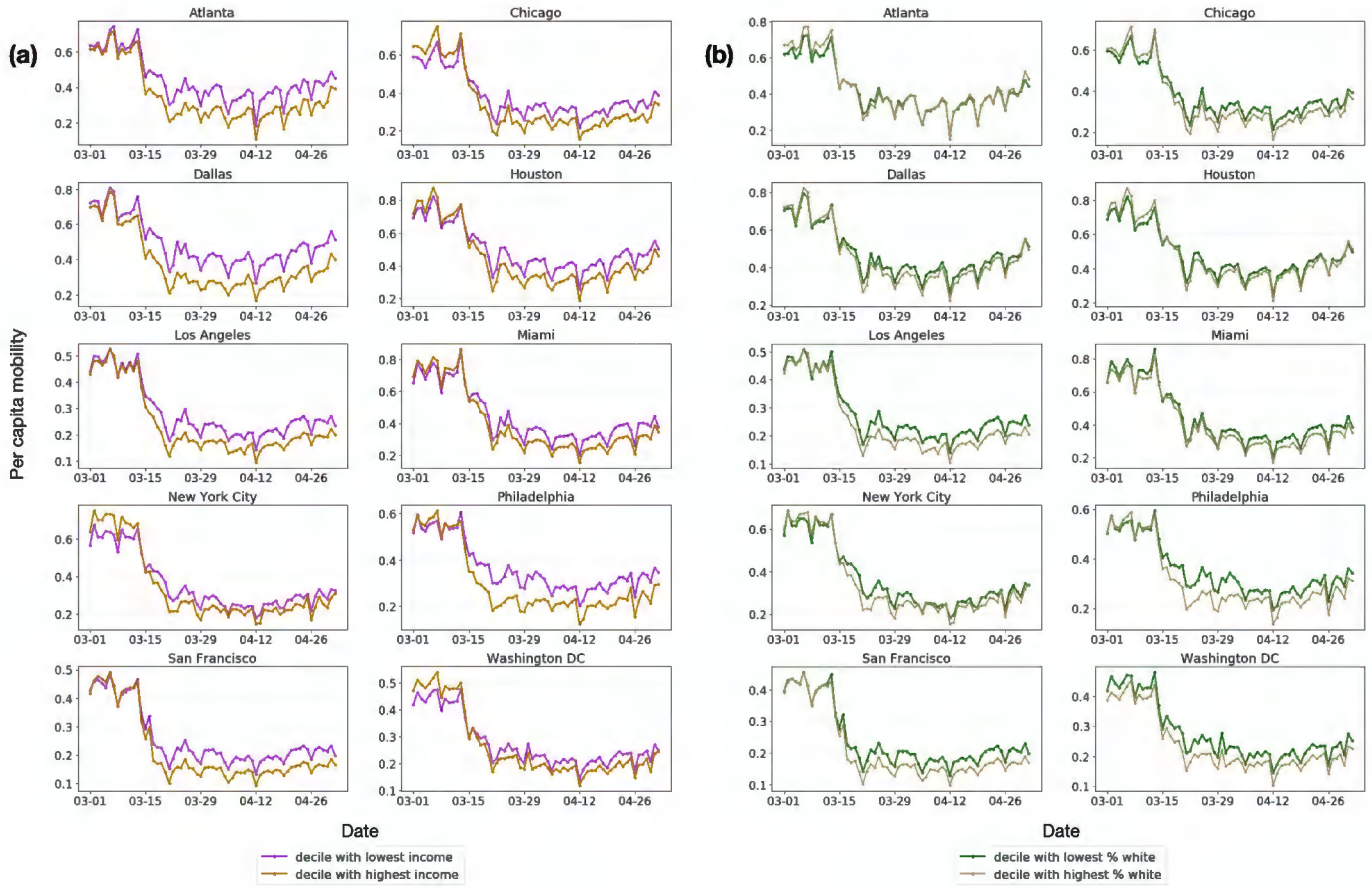
Extended Data Fig. 3 | Trade-off between new infections and visits lost from reopening. We simulate reduced maximum occupancy reopening starting on 1 May 2020 and run the simulation until the end of the month. Each dot represents the level of occupancy reduction: for example, capping visits at 50% of the maximum occupancy. The *y* coordinate represents the predicted number of new infections incurred after reopening (per 100,000 population) and the *x* coordinate represents the fraction of visits lost from partial reopening compared to full reopening. Shaded regions denote the 2.5th and 97.5th percentiles across parameter sets and stochastic realizations. In four metro areas, the predicted cost of new infections from reopening is roughly similar for lower-income CBGs and the overall population, but in five metro areas, the lower-income CBGs incur more predicted infections from reopening. Notably, New York City (NYC) is the only metro area in which this trend is reversed; this is because the model predicts that such a high fraction—65% (95% confidence interval, 62–68%)—of lower-income CBGs in NYC had been infected before reopening that after reopening, only a minority of the lower-income population is still susceptible (in comparison, the second highest fraction infected before reopening was 31% (95% confidence interval, 28–35%) for Philadelphia, and the rest ranged from 1 to 14%).



Extended Data Fig. 4 | Reduced maximum occupancy versus uniform reduction reopening. In comparison to partially reopening by uniformly reducing visits, the reduced maximum occupancy strategy—which disproportionately targets POIs during their most risky high-density periods—always results in a smaller predicted increase in infections for the same number of visits. The *y* axis plots the relative difference between the predicted increase in cumulative infections (from 1 May to 30 May 2020) under the reduced occupancy strategy compared to the uniform reduction strategy. The shaded regions denote the 2.5th and 97.5th percentiles across the sampled parameters and stochastic realizations.



Extended Data Fig. 5 | POI attributes in all 10 metro areas combined. a, **b**, The POIs from all metro areas are pooled and the quantities from the mobility data are shown. **a**, The distribution of dwell time. **b**, The average number of hourly visitors divided by the area of the POI in square feet. Each point represents one POI; boxes depict the interquartile range across POIs; data points outside the range are shown as individual dots. **c**, **d**, The data are pooled across model realizations from all metro areas and model predictions are shown for the increase in infections (per 100,000 population) because of reopening a POI category. **c**, Data per POI. **d**, Data for the category as a whole. Each point represents a model realization; boxes depict the interquartile range across realizations; data points outside the range are shown as individual dots. Across MSAs, we model 552,758 POIs in total, and we sample 97 parameters and 30 stochastic realizations (n = 2,910); see Supplementary Table 6 for the number of sets per metro area. Colours are used to distinguish the different POI categories, but do not have any additional meaning.



Extended Data Fig. 6 | Daily per-capita mobility over time. a, **b**, We compare mobility in the lowest and highest deciles of CBGs based on median household income (**a**) and the percentage of white residents (**b**). See Methods, 'Analysis details' for details.

Extended Data Table 1 | Dataset summary statistics from 1 March to 2 May 2020

Metro area	CBGs	POIs	Hourly edges	Total modeled pop	Total visits
Atlanta	3,130	39,411	540,166,727	7,455,619	27,669,692
Chicago	6,812	62,420	540,112,026	10,169,539	33,785,702
Dallas	4,877	52,999	752,998,455	9,353,561	37,298,053
Houston	3,345	49,622	609,766,288	7,621,541	32,943,613
Los Angeles	8,904	83,954	643,758,979	16,101,274	38,101,674
Miami	3,555	40,964	487,544,190	6,833,129	26,347,947
New York City	14,763	122,428	1,057,789,207	20,729,481	66,581,080
Philadelphia	4,565	37,951	304,697,220	6,759,058	19,551,138
San Francisco	2,943	28,713	161,575,167	5,137,800	10,728,090
Washington DC	4,051	34,296	312,620,619	7,740,276	17,898,324
All metro areas combined	**56,945**	**552,758**	**5,411,028,878**	**97,901,278**	**310,905,313**

Extended Data Table 2 | Model parameters

Param.	Description	Value (Source)
δ_E	mean latency period	96 hours[24,63]
δ_I	mean infectious period	84 hours[24]
δ_c	period from infectious to confirmed	7 days[24,71]
r_c	percentage of cases which are detected	10%[24,63,69–71]
β_{base}	base CBG transmission rate	Variable (Estimated)
N_{c_i}	population size of CBG c_i	Variable (2018 US Census[57])
ψ	scaling factor for POI transmission	Variable (Estimated)
$w_{ij}^{(t)}$	# visitors from CBG c_i to POI p_j at time t	Variable (SafeGraph)
a_{p_j}	area of POI p_j in square feet	Variable (SafeGraph)
p_0	initial proportion of exposed population	Variable (Estimated)
$S_{c_i}^{(0)}$	initial susceptible population in CBG c_i	$(1-p_0)N_{c_i}$
$E_{c_i}^{(0)}$	initial exposed population in CBG c_i	$p_0 N_{c_i}$
$I_{c_i}^{(0)}$	initial infectious population in CBG c_i	0
$R_{c_i}^{(0)}$	initial removed population in CBG c_i	0

If the parameter has a fixed value, we specify it under 'Value'; otherwise, 'Variable' is used to indicate that it varies across CBG, POI or metro area.

Extended Data Table 3 | Predicted transmission rate disparities at each POI category between income groups

Metro area	ATL	CHI	DAL	HOU	LA	MIA	NY	PHL	SF	DC	**Median**
Full-Service Restaurants	0.764	1.204	0.956	1.000	1.445	1.232	2.035	2.883	1.758	1.171	1.218
Limited-Service Restaurants	0.940	0.950	1.002	0.906	1.067	0.872	1.901	1.614	0.994	0.962	0.978
Other General Stores	0.782	1.083	0.957	0.729	0.760	0.894	1.218	1.312	1.045	0.950	0.954
Gas Stations	1.326	1.865	1.310	1.515	2.254	2.195	1.899	6.461	1.357	1.870	1.868
Fitness Centers	0.536	0.907	0.708	0.670	1.461	0.789	1.151	1.516	0.995	1.160	0.951
Grocery Stores	0.948	3.080	0.838	1.333	2.408	1.498	4.984	10.437	2.478	1.977	2.192
Cafes & Snack Bars	1.385	0.919	0.716	1.120	1.327	2.168	1.943	1.757	0.982	0.932	1.224
Hotels & Motels	1.228	1.200	0.814	0.804	1.229	1.134	1.260	1.993	1.199	1.346	1.214
Religious Organizations	1.546	1.763	0.956	0.919	1.746	1.464	1.756	1.736	1.515	1.852	1.641
Hardware Stores	3.938	3.340	1.575	2.111	1.333	0.939	3.553	6.716	4.202	13.560	3.446
Department Stores	1.132	1.230	0.978	0.911	1.083	1.431	1.667	0.976	0.867	1.042	1.062
Offices of Physicians	1.235	0.721	0.667	1.036	1.141	1.687	1.307	1.319	1.193	0.445	1.167
Pharmacies & Drug Stores	1.636	1.389	1.176	0.854	1.718	1.555	2.577	5.624	1.200	1.699	1.596
Sporting Goods Stores	0.936	1.540	1.129	0.812	1.168	0.700	1.253	1.161	0.826	2.777	1.145
Automotive Parts Stores	0.890	1.707	0.862	1.086	1.990	1.414	1.524	2.697	1.753	1.246	1.469
Used Merchandise Stores	0.993	0.931	1.000	1.315	1.017	1.074	1.352	1.668	1.587	0.814	1.046
Convenience Stores	1.208	0.932	1.613	0.647	0.838	0.824	1.736	2.322	1.086	1.428	1.147
Pet Stores	1.260	0.820	1.192	1.487	1.536	0.776	3.558	1.652	2.124	0.905	1.374
New Car Dealers	2.036	1.471	0.741	0.809	1.180	1.377	2.022	1.129	0.395	0.872	1.154
Hobby & Toy Stores	1.168	1.110	1.165	0.853	1.771	1.520	1.525	1.088	0.883	0.926	1.138
Median	1.188	1.202	0.968	0.915	1.330	1.305	1.746	1.702	1.196	1.166	

We report the ratio of the average predicted transmission rate encountered by visitors from CBGs in the bottom income decile to that for the top income decile. A ratio greater than 1 means that visitors from CBGs in the bottom income decile experienced higher (more dangerous) predicted transmission rates. See Methods, 'Analysis details' for details.

Extended Data Table 4 | Predicted transmission rate disparities at each POI category between racial groups

Metro area	ATL	CHI	DAL	HOU	LA	MIA	NY	PHL	SF	DC	**Median**
Full-Service Restaurants	0.802	1.354	0.981	0.965	1.065	1.167	2.418	2.661	1.223	1.013	1.116
Limited-Service Restaurants	0.940	1.144	1.028	0.940	0.820	0.919	2.136	1.523	0.799	1.346	0.984
Other General Stores	0.776	1.277	0.838	0.841	1.527	1.132	2.158	1.313	0.925	1.312	1.204
Gas Stations	1.402	1.891	1.389	1.190	1.336	1.857	1.818	2.286	2.321	1.316	1.610
Fitness Centers	0.607	1.167	0.670	0.831	0.780	1.066	1.447	1.977	1.103	1.205	1.084
Grocery Stores	0.589	3.664	0.613	1.195	2.386	0.950	5.864	13.705	2.243	2.262	2.252
Cafes & Snack Bars	1.308	1.104	0.845	0.840	0.976	2.619	1.767	2.456	1.045	0.867	1.074
Hotels & Motels	0.977	1.007	1.366	0.718	1.112	1.024	1.449	2.494	0.654	0.899	1.015
Religious Organizations	0.938	1.606	1.060	0.953	2.096	1.795	1.933	2.040	1.674	1.188	1.640
Hardware Stores	0.909	3.900	1.523	1.461	1.952	0.586	5.032	3.898	11.103	13.432	2.925
Department Stores	1.081	1.301	0.805	0.777	0.992	2.337	2.479	1.357	1.089	1.402	1.195
Offices of Physicians	0.894	1.323	1.006	1.415	0.898	1.117	1.652	2.073	0.694	1.911	1.220
Pharmacies & Drug Stores	0.888	1.376	0.930	0.732	1.538	1.674	3.315	3.366	1.135	1.715	1.457
Sporting Goods Stores	0.767	0.674	0.650	0.506	1.946	0.818	1.532	2.152	0.880	1.715	0.849
Automotive Parts Stores	1.049	1.479	1.010	1.353	2.998	2.657	1.740	3.387	1.646	0.601	1.562
Used Merchandise Stores	0.858	1.195	0.699	1.060	1.270	0.593	1.500	3.024	1.425	0.799	1.128
Convenience Stores	2.016	5.055	1.272	2.188	0.761	0.902	1.911	2.276	1.239	1.844	1.878
Pet Stores	0.925	1.624	0.724	1.465	1.506	0.881	2.715	10.182	1.568	2.408	1.537
New Car Dealers	1.008	1.398	0.812	0.736	0.942	0.998	1.977	0.866	0.772	0.383	0.904
Hobby & Toy Stores	2.569	0.853	0.628	0.979	1.373	1.388	2.237	0.825	0.864	1.286	1.132
Median	0.932	1.339	0.888	0.959	1.303	1.092	1.955	2.281	1.119	1.314	

We report the ratio of the average predicted transmission rate encountered by visitors from CBGs with the lowest (bottom decile) proportion of white residents versus that for the top decile. A ratio greater than 1 means that visitors from CBGs in the bottom decile experienced higher (more dangerous) predicted transmission rates. See Methods, 'Analysis details' for details.

nature research

Corresponding author(s): Jure Leskovec

Last updated by author(s): Oct 7, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No data collection was performed for this study; all analysis relied on previously collected datasets, as described in the Data section below.
Data analysis	All data analysis was performed using Python with standard libraries. Code is available at http://snap.stanford.edu/covid-mobility and https://github.com/snap-stanford/covid-mobility.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Census data (https://www.census.gov/programs-surveys/acs), case and death counts from The New York Times (https://github.com/nytimes/covid-19-data), and Google mobility data (https://www.google.com/covid19/mobility/) are publicly available. Cell phone mobility data is freely available to researchers, non-profits, and governments through the SafeGraph COVID-19 Data Consortium (https://www.safegraph.com/covid-19-data-consortium).

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☒ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	This is a quantitative epidemiological modeling study.
Research sample	We study previously-collected data on 10 of the largest American metropolitan statistical areas. The data came from the following: 1) Anonymized mobility data from smartphone users from SafeGraph. SafeGraph data is obtained by "partnering with mobile applications that obtain opt-in consent from its users to collect anonymous location data", per official SafeGraph documentation. 2) US Census 3) Case and death counts from The New York Times 4) Google mobility data
Sampling strategy	We did not perform sampling, but relied on previously collected datasets. All datasets were chosen as they were comprehensive: the case and death counts and Census data were designed to cover the entire population; Google mobility data covers a large subset of the population; and previous analyses have shown that the SafeGraph anonymized cell phone mobility dataset is geographically representative: for example, it does not systematically over-represent individuals from higher-income areas (https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset).
Data collection	We did not perform data collection, but relied on previously collected datasets. Census data is collected as described here: https://www.census.gov/programs-surveys/acs. The cell phone mobility data is collected from cell phone geolocation information, and is anonymized and aggregated. We accessed and downloaded the datasets in May 2020.
Timing	We make use of data in the following ranges: 1) Safegraph mobility data: Jan 1, 2019 - May 2, 2020 2) US Census: 2013-2018 3) Case and death counts from The New York Times: March 8 - May 9, 2020 4) Google mobility data: March 1 - May 2, 2020
Data exclusions	The original SafeGraph dataset contains 5.4 million points of interest (POIs). We retain a POI in our final dataset if it satisfies the following criteria: (1) it lies within one of the 10 American metropolitan areas that we analyze (out of 384 metropolitan statistical areas total); (2) SafeGraph has visit data for this POI for every hour that we model, from 12am on March 1, 2020 to 11pm on May 2, 2020; (3) SafeGraph has recorded the home CBGs of this POI's visitors for at least one month from January 2019 to February 2020; (4) the POI is not a "parent" POI, as defined in the Methods section. After applying these filters, our dataset contains 553k POIs. Most POIs are filtered out because they do not lie within the 10 large metropolitan statistical areas that we study; this filtering decision was made prior to any analysis for computational tractability reasons. In our analysis of POI-specific category risks, we do not analyze 6 categories of POIs because we wish to be conservative and only focus on categories where we are most confident we are fully capturing transmission at the category: Child Day Care Services, Elementary and Secondary Schools, Drinking Places (Alcoholic Beverages), Nature Parks and Other Similar Institutions, General Medical and Surgical Hospitals, and Other Airport Operations. The justifications for these exclusions, which are based on prior work, are given in the Methods section.
Non-participation	Because we relied on previously collected anonymized, aggregated data from cell phone mobility tracking, we did not have access to individual-level data and do not know how many participants dropped out/declined participation.
Randomization	This is not a randomized controlled trial and participants were not randomized into experimental groups.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☒	☐ Clinical data
☒	☐ Dual use research of concern

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	See above.
Recruitment	See above. Prior work has looked into biases in the SafeGraph dataset (https://www.safegraph.com/blog/what-about-bias-in-the-safegraph-dataset).
Ethics oversight	The dataset from The New York Times consists of aggregated COVID-19 confirmed case and death counts collected by journalists from public news conferences and public data releases. For the mobility data, consent was obtained by the third-party sources collecting the data. SafeGraph aggregates data from mobile applications that obtain opt-in consent from their users to collect anonymous location data. Google's mobility data consists of aggregated, anonymized sets of data from users who have chosen to turn on the Location History setting. Additionally, we obtained IRB exemption for SafeGraph data from the Northwestern University IRB office.

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Exhibit D

J Korean Med Sci. 2020 Nov 30;35(46):e415
https://doi.org/10.3346/jkms.2020.35.e415
eISSN 1598-6357·pISSN 1011-8934

JKMS

Original Article
Preventive & Social Medicine

Evidence of Long-Distance Droplet Transmission of SARS-CoV-2 by Direct Air Flow in a Restaurant in Korea

Keun-Sang Kwon [1,2*] **Jung-Im Park** [2*] **Young Joon Park** [3] **Don-Myung Jung** [4] **Ki-Wahn Ryu** [5] **and Ju-Hyung Lee** [1,2]

[1]Department of Preventive Medicine, Jeonbuk National University Medical School, Jeonju, Korea
[2]Jeonbuk Center for Infectious Disease Control and Prevention, Jeonju, Korea
[3]Korea Disease Control and Prevention Agency, Cheongju, Korea
[4]Division of Health Care, Jeonbuk Provincial Government, Jeonju, Korea
[5]Department of Aerospace Engineering, College of Engineering, Jeonbuk National University, Jeonju, Korea

Received: Oct 6, 2020
Accepted: Nov 16, 2020

Address for Correspondence:
Ju-Hyung Lee, MD, PhD
Department of Preventive Medicine, Jeonbuk National University Medical School, 20 Geonji-ro, Deokjin-gu, Jeonju 54907, Republic of Korea.
E-mail: premd77@jbnu.ac.kr

*Keun-Sang Kwon and Jung-Im Park contributed equally to this work.

ORCID iDs
Keun-Sang Kwon
https://orcid.org/0000-0001-6795-6881
Jung-Im Park
https://orcid.org/0000-0002-7831-3694
Young Joon Park
https://orcid.org/0000-0002-5971-7829
Don-Myung Jung
https://orcid.org/0000-0002-9301-1897
Ki-Wahn Ryu
https://orcid.org/0000-0002-4575-4493
Ju-Hyung Lee
https://orcid.org/0000-0003-2487-4098

Disclosure
The authors have no potential conflicts of interest to disclose.

ABSTRACT

Background: The transmission mode of severe acute respiratory syndrome coronavirus 2 is primarily known as droplet transmission. However, a recent argument has emerged about the possibility of airborne transmission. On June 17, there was a coronavirus disease 2019 (COVID-19) outbreak in Korea associated with long distance droplet transmission.
Methods: The epidemiological investigation was implemented based on personal interviews and data collection on closed-circuit television images, and cell phone location data. The epidemic investigation support system developed by the Korea Disease Control and Prevention Agency was used for contact tracing. At the restaurant considered the site of exposure, air flow direction and velocity, distances between cases, and movement of visitors were investigated.
Results: A total of 3 cases were identified in this outbreak, and maximum air flow velocity of 1.2 m/s was measured between the infector and infectee in a restaurant equipped with ceiling-type air conditioners. The index case was infected at a 6.5 m away from the infector and 5 minutes exposure without any direct or indirect contact.
Conclusion: Droplet transmission can occur at a distance greater than 2 m if there is direct air flow from an infected person. Therefore, updated guidelines involving prevention, contact tracing, and quarantine for COVID-19 are required for control of this highly contagious disease.

Keywords: Infectious Disease Transmission; SARS-CoV-2; COVID-19

INTRODUCTION

The coronavirus disease 2019 (COVID-19) caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) had spread throughout the world, and the total cases numbered more than 17 million and 680,000 deaths by COVID-19 as of August 3, 2020.[1] On January 20, 2020, a Chinese traveler from Wuhan, China, was identified as the first COVID-19 case in Korea, after which only 30 cases of COVID-19 were reported until February 20 and were based on visitors to the country and contact with them. After an outbreak associated with a religious group in Daegu Metropolitan city, the number of new patients per day rapidly increased to a maximum of 813 from late February to late March. The government of Republic of Korea had increased the response level for public health emergency to level 3 (from level

Author Contributions

Conceptualization: Kwon KS, Park YJ, Lee JH; Investigation: Park JI, Jung DM, Ryu KW, Lee JH; Data curation: Park JI; Writing - original draft: Kwon KS, Park JI; Writing - review & editing: Kwon KS, Park JI, Park YJ, Jung DM, Ryu KW, Lee JH.

0 to level 3) on February 23, and had implemented high-intensity social distancing until May 5.[2] Furthermore, the government applied the K-quarantine model '3T policy (Test-Trace-Treat)' system that included a rapid and exact test for COVID-19, investigation of the epidemic using information and communication technology (ICT), and an isolation and care program according to severity.[3]

On June 17, there was a new confirmed COVID-19 case (index case, case A) in Jeonju, Korea, considered as transmitted by droplets at 6.5 m away from the infector and 5 minutes exposure in a restaurant with air conditioning. It is important to know how SARS-CoV-2 is transmitted between people in various situations. We share these investigation results as a reference to update guidelines involving prevention, tracing, and quarantine for control of this pandemic infectious disease.

METHODS

Personal factor investigation

The epidemiological investigation was implemented according to the 'Infectious Diseases Control and Prevention Act' (Act number 16725) in Korea and the guidelines for response to COVID-19 by the Korea Disease Control and Prevention Agency (KDCA).[4] Data comprising patient's personal statements by interview, medical institution usage history, credit card record, closed-circuit television (CCTV) images, cell phone location data, and other associated information were secured by an epidemiological investigation team.[5] The Epidemic Investigation Support System (EISS) developed by KDCA was also used for location tracking of confirmed cases and hot-spot analysis.[6] Nasopharyngeal specimens of cases and close contacts were collected and tested using real-time reverse transcription polymerase chain reaction (rRT-PCR) by Jeollabuk-do Institute of Health & Environment Research, and genome sequencing analysis for verifying association between cases was performed by the KDCA.[7] In total, 10–100 ng of the extracted viral RNA with a maximum volume of 8.5 μL was subjected to target enrichment using a Truseq RNA library prep for enrichment (Illumina, San Diego, CA, USA) and Truseq RNA Enrichment (Illumina). Dual-index filtering and adapter trimming were conducted on the sequences using our in-house scripts. Hybridization probes were designed to cover the whole genome of SARS-CoV-2 using the Wuhan-Hu-1 strain.[8] The biotinylated probes were 120 base pair in length with 3 × tiling (Celemics, Inc. Seoul, Korea). In total, 745 conserved probes were generated.

Environmental factor investigation

As a result of EISS analysis, one restaurant (restaurant A) was visited by case A on June 12 and was identified as the site of exposure. The first field investigation was started from June 19 based on assessment of CCTV, table locations, timeline, and movement route of case A and other people in the restaurant were verified. Also, the internal structure, distance between visitors, and exact locations of the ceiling type air conditioners were investigated. Air speed and direction at several specified positions were precisely measured using a portable anemometer (Kestrel 2500; Nielsen-Kellerman Co. Boothwyn, PA, USA) on June 24 and July 2. To measure air flow, we set the air conditioner at the same fan speed and direction of June 12. The chairs of cases and visitors were also occupied by people to simulate the same situation. A total of 39 environmental samples of inlets and outlets of air conditioners, table seat of case A, and nearby tables and chairs in consideration of air flow direction were collected on June 23 for testing of SARS-CoV-2 in the environment and were analyzed by rRT-PCR test.[9]

https://doi.org/10.3346/jkms.2020.35.e415

RESULTS

Personal factors associated with the outbreak

The symptoms of the index case (case A) started on June 16, and the probable period of exposure was assumed to be from June 2 to June 15 according to the incubation period of SARS-CoV-2. Because case A had no history of overseas travel and travel outside Jeonju, where there had been no confirmed case in the previous 2 weeks, we used the EISS of the KDCA to gather data from June 2 to June 15. The results showed only one (case B) of 538 confirmed domestic cases with a tracking map overlapping that of case A during that period. The site of overlap was a restaurant (restaurant A), where case A and case B were co-located for 5 minutes on June 12 based on CCTV images. Case B lives in Daejeon Metropolitan city within an hour's drive away from Jeonju, and visited Jeonju only on June 12. Therefore, we tentatively considered case B is the primary case (infector) and case A as the secondary case (infectee). When case B was in the restaurant, they came into close contact with 11 visitors and two employees who did not properly wear a mask. The epidemiological investigation team ordered these contacts to undergo the rRT-PCR test for SARS-CoV-2 on June 19 or 20, and to perform a minimum 14-day quarantine until June 26. Among the 13 close contacts during the quarantine period, one additional case (case C, visitor) was confirmed to have COVID-19 on June 20. Finally, case B spread COVID-19 to two infectees in this restaurant, for an attack rate of 15.4% (2/13). The exposure day was June 12, and the symptom onset of case B was June 13, resulting in a median incubation period and serial interval of 5 days and 4 days, respectively, consistent with previous reports.[10-12] There was another case in that restaurant (case D). Case D is a companion of case B and visited this restaurant with case B on June 12. Case D had been exposed to another epidemic patient in Daejeon city on June 11 and showed a positive result for SARS-CoV-2 on June 16, with symptom onset the day before. Because case D could not spread COVID-19 to other people at that time, case D was excluded from this outbreak investigation (**Figs. 1** and **2**).

Environmental factors and mode of transmission

Restaurant A was located on the first floor of a six-story building totaling 96.6 square meters in size (9.2 × 10.5 m) without windows or a ventilation system. There were two doors in the restaurant, one at the front (door 1) and the other at the back (door 2). Two ceiling-type air conditioners were diagonally located at 3.2 m from the floor as shown in **Fig. 3**; they were fixed with wire and had been operating at the time the cases were in restaurant A. On CCTV, case A and his companion entered the restaurant at 16:00 on June 12 and finished their meals before case B (with case D) entered at 17:15 using door 2. Case B and his colleague sat at a table near

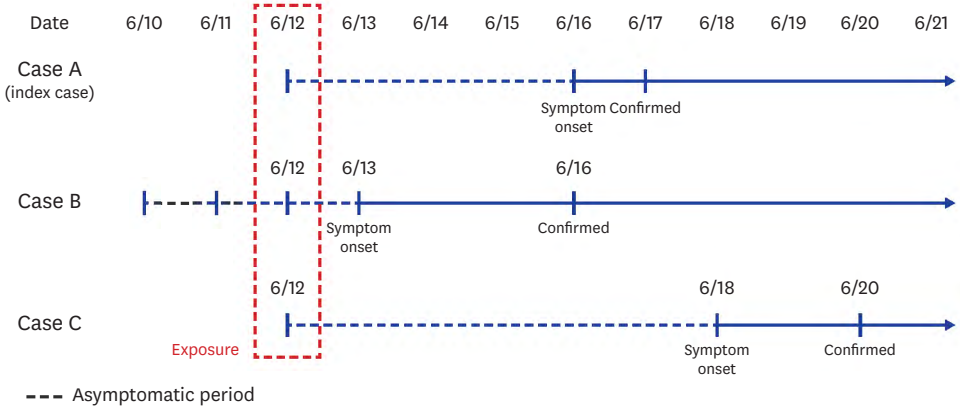


Fig. 1. The asymptomatic period and symptom onset of all three coronavirus disease 2019 cases.

https://doi.org/10.3346/jkms.2020.35.e415

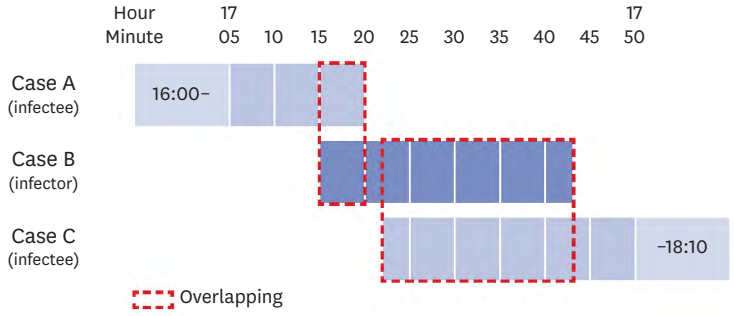


Fig. 2. Timeline of coronavirus disease 2019 infector and infectees in the restaurant. Case A (index case) overlaps about 5 minutes (17:15–17:20) with case B (infector), who overlaps case C for about 21 minutes (17:22–17:43).

door 2, at a 6.5-m distance from case A, who did not leave from his table or share his table with others. Cases A and B engaged in conversation with their respective companions without masks. At 17:20, case A went out of the restaurant A using door 1. In 2 minutes, case C and his companions (V6, V7) entered the restaurant A using door 1 and sat at another table 4.8 m distant from case B, where they remained for 21 minutes before case B left his table using door 1 at 17:43. The distance between case A and case B was 6.5 m, and the air flow direction at positions of both cases showed a maximum of 1.0 m/sec (3.6 km/hr) velocity measured by anemometer. The air flow between case B and case C showed a maximum of 1.2 m/sec (4.3 km/hr) over a 4.8 m distance. All positions such as guest tables, infectors, and infectees, ceiling air conditioners, and information for air speed and direction are shown in **Fig. 3**.

The 39 environment samples for SARS-CoV-2 were all negative by rRT-PCR. The results of genome sequencing of the three patients were presented as all three cases' genomic types were GH type and identical for each other (data not shown).

Summary of epidemiological findings and implications: 1) Indoor air flow (maximum velocity, 1.0 m/sec) could have transmitted droplets from the infector (case B) to infectee (case A) within 6.5 m and 5 minutes of exposure and to a second infectee (case C) within 4.8 m and 21 minutes of exposure. 2) The attack rate among exposures at the restaurant was 15.4% (2/13; 95% CI, 8.3%–22.5%) and is higher than the secondary attack rate among the total close contacts (0.6%; 0.3%–1.0%) and only household contacts (7.6%; 3.7%–14.3%) but was similar to a call center exposure outbreak (15.1%; 10.8%–20.6%).[13,14] 3) COVID-19 transmission by droplets from an infector can occur over a greater than 2 m distance with a short period of exposure when combined with air flow. The guidelines on quarantine and epidemiological investigation must be updated to reflect these factors for control and prevention of COVID-19. 4) The EISS for tracing COVID-19 patients was very useful and shortened the time to locate the infection source. This system could suppress the regional epidemic scale of the virus and reduce the burden for testing, isolation, and treatment.

DISCUSSION

The SARS-CoV-2 virus is mainly transmitted through respiratory droplets emitted from an infector's coughing, sneezing, talking, and normal breathing and upon close contact between people.[15] These droplets are typically divided into large and small sizes (also called aerosol) based on a diameter of 5 μm. The large particles (also called droplets) tend to settle within 1–2 m of their origin due to gravitational force, and the settling velocity is proportional

https://doi.org/10.3346/jkms.2020.35.e415

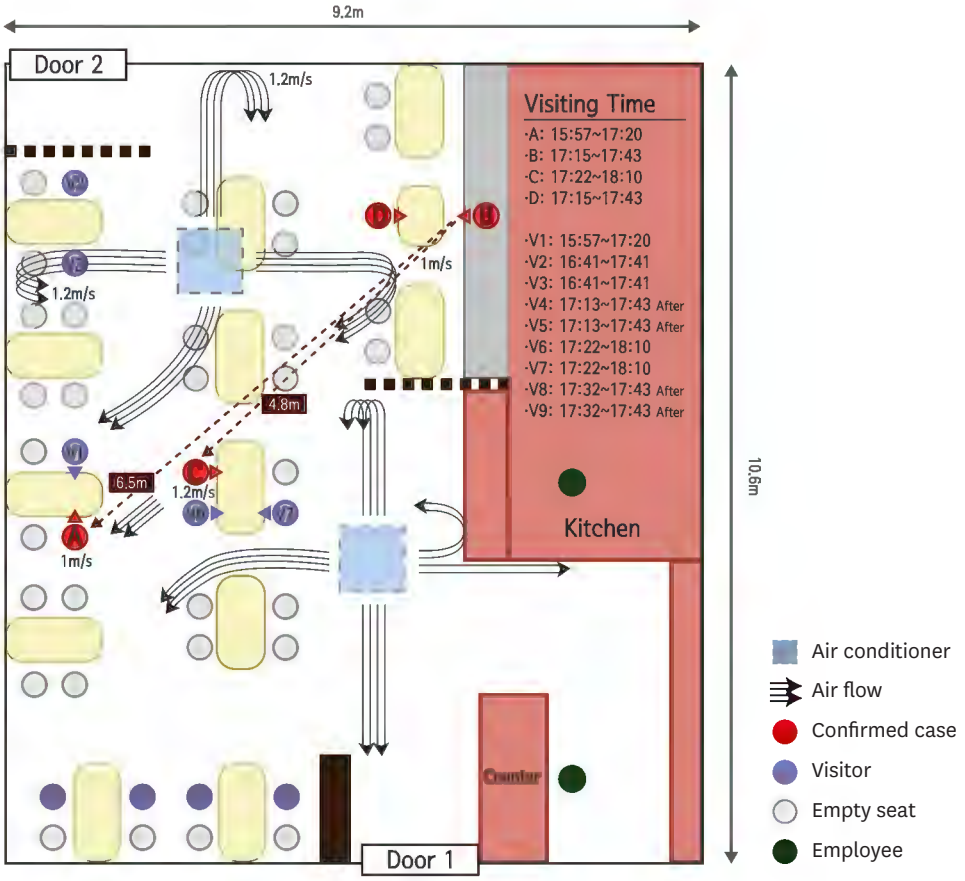


Fig. 3. Schematic diagram of the outbreak restaurant equipped with ceiling-type air conditioners. The arrowed solid streamlines represent the air flow directions in the restaurant. Curved air streamlines represent that air streams from the ceiling air conditioners are reflected from the wall or barrier, and move downward toward the floor.

to the particle diameter.[16] Therefore, social distancing requires a minimum of 1–2 m to avoid contact with a virus-containing respiratory droplet. In the situation of no effective treatment drug or vaccine, the most important personal methods to prevent or control this pandemic of COVID-19 are social distancing, use of a mask (if social distancing cannot be maintained), and handwashing.

Recently, it has been suggested that COVID-19 can be spread through not only droplets or contact, but also airborne transmission. An experimental study showed that the COVID-19 virus in aerosol particles remained viable during 3 hours and 16 hours.[17] Morawska and Milton, together with their co-authors, 239 scientists, strongly suggested the possibility of airborne transmission of COVID-19 based on several preprint findings, though there has been no peer review of this research.[18] The last updated version of the WHO scientific brief reported on July 9, 2020 reported that airborne transmission by aerosols is rare, and SARS-CoV-2 is spread primary between people through droplets or close contact. However, the possibility of aerosol transmission in crowded indoor spaces has been suggested in combination with droplet transmission.[19]

In this outbreak, the distances between infector and infected persons were 4.8 and 6.5 m, both farther than the generally accepted 2 m droplet transmission range. This is some of the first evidence of airborne transmission. At the field investigation, we assumed the possibility

https://doi.org/10.3346/jkms.2020.35.e415

of long-distance droplet movement by air flow. Dbouk and Drikakis reported results of computational fluid dynamics showing that most droplets settle within 1–2 m in the absence of airflow. However, with a 4 km/hr or 15 km/hr wind, droplets could travel 6 m after 5 or 1.6 seconds, respectively.[20] In the presented case, the air flow from infector to infectee showed a 1.0–1.2 m/sec (3.6–4.3 km/hr) velocity, indicating the need for 6.5 seconds of contact to transmit droplets from the infector to the index case. Only the visitors (cases A and C) sitting in the air flow path of case B were infected with COVID-19, while other visitors (V2, V3) closer to the infector for a longer period of time but in the absence of direct air flow did not become infected. In addition, the visitors sitting at tables with cases A and C (V1, V6, and V7) were not infected with COVID-19 because they faced away from the infector's face. These findings strongly suggest that this outbreak occurred by droplet transmission exceeding a 2 m distance and excluded contact and fomite transmission. This transmission pattern is similar with the outbreak of a restaurant with air conditioning in Guangzhou, China.[21] In this article, the authors concluded that the most likely transmission was done by droplet and also emphasized the direction of air flow.

Without the K-quarantine model (test-trace-treat) and EISS, it may have been very difficult to establish an infection chain of this outbreak because the incubation period of COVID-19 has a wide range from 1 to 14 days, and the exposure occurred only once for 5 minutes at a 6.5 m distance. Based on this system, the infector was identified within 2 days after the index case was confirmed. This short period of identification could simplify identification of close contacts and reduce outbreak size by quarantining of all close contacts. In most COVID-19 outbreak situations, identification of the infection chain is difficult or almost impossible. An outbreak for which an infectious source cannot be determined may be due to failure of consideration of a short period of exposure or of clearly verifying the movements of the confirmed person. In addition, it suggests that some suspected airborne transmission reports may be misinterpreted based on lack of awareness of the long transmission mechanism of droplets.

The guideline for control of droplet transmission over a long distance (above 2 m) by air flow in indoor settings are similar to those of airborne transmission and comprise sufficient ventilation and social distancing (avoid overcrowding, maintain distance between people).[22] However, if there is high possibility of transmission by aerosol or droplet transmission over a long distance, N95 respiratory or equivalent masks are needed not only in health care settings, but in any indoor environment. In Korea, the Ministry of Food and Drug Safety is concerned with approval of medical mask like KF99, KF94, and KF80 and developed a type of mask named KF-AD (anti-droplet) for COVID-19. Any such mask, including KF and surgical, that can protect against droplets should be sufficient for preventing droplet transmission.

According to this case, additional considerations need to occur for COVID-19 prevention and control. The first is that transmission in an indoor setting is possible at a distance greater than 2 m with a short period of exposure (five minutes), and selection of close contacts in contact tracing should be changed. When epidemiological field investigation of an indoor environment is needed, it is necessary to assess the seating arrangement and operation and location of fans (including ceiling fans) or air conditioners with wind direction and velocity. It is also necessary to ventilate frequently for management of indoor air or to apply a ventilation system or forced ventilation method if natural ventilation is not possible. Furthermore, the distance between tables at an indoor restaurant or cafeteria should be greater than 1–2 m, or installation of a wind partition should be considered based on air flow. In addition, in indoor

settings such as restaurants, masks should be removed only during meals and should be worn before and after eating, while conversation during meals and loud talking or shouting should be avoided. In the long term, installation of separate rooms or bulkheads for indoor settings should be considered to prevent transmission of airborne and droplet infectious diseases.

There are some limitations to this study. First, we did not assess air flow using computational fluid dynamics. In addition, the air flow measurement can't reflect all the same situations because opening of doors and motions of cases and visitors were not reproduced. However, air flow direction and velocity were identified between the infector and infectees using an anemometer in the most similar environment as possible. Second, environmental samples were collected at 11 days after the inspector visit. Though all these results were negative, this was not proof against airborne transmission.

In conclusion, droplet transmission can occur at a distance greater than 2 m if there is direct air flow from an infected person in an indoor setting. Therefore, updated guidelines for quarantine and environmental management of COVID-19 are needed until approval of an effective treatment drug or vaccine.

ACKNOWLEDGMENTS

The authors appreciate the epidemiological investigation team of Jeonju and Jeollabuk-do and all researchers in Jeonbuk Center for Infectious Disease Control and Prevention. Also, we thank all members of the Department of Infectious Disease of Jeollabuk-do Institute of Health and Environment.

REFERENCES

1. World Health Organization. Coronavirus disease 2019 (COVID-19): situation report-195. https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200802-covid-19-sitrep-195.pdf. Updated 2020. Accessed July 30, 2020.
2. Park K, Kim Y, Yeom H, Hwang I, Kwon J, Kim M, et al. Weekly report on the COVID-19 situation in the Republic of Korea (as of July 25, 2020). *Public Health Wkly Rep* 2020;13(31):2264-78.
3. Ministry of Economy and Finance. Flattening the curve on COVID-19 - how Korea responded to a pandemic using ICT. http://english.moef.go.kr/pc/selectTbPressCenterDtl.do?boardCd=N0001&seq=4879. Updated 2020. Accessed July 30, 2020.
4. National Law Information Center. Infectious Diseases Control and Prevention Act [Act number 16725]. http://www.law.go.kr/LSW/eng/engLsSc.do?menuId=2&section=lawNm&query=infectious&x=0&y=0#liBgcolor5. Updated 2020. Accessed July 30, 2020.
5. COVID-19 National Emergency Response Center, Epidemiology & Case Management Team, Korea Centers for Disease Control & Prevention. Contact transmission of COVID-19 in South Korea: novel investigation techniques for tracing contacts. *Osong Public Health Res Perspect* 2020;11(1):60-3.
PUBMED | CROSSREF
6. Park YJ, Cho SY, Lee J, Lee I, Park WH, Jeong S, et al. Development and utilization of a rapid and accurate epidemic investigation support system for COVID-19. *Osong Public Health Res Perspect* 2020;11(3):118-27.
PUBMED | CROSSREF
7. Corman VM, Landt O, Kaiser M, Molenkamp R, Meijer A, Chu DK, et al. Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Euro Surveill* 2020;25(3):2000045.
PUBMED | CROSSREF
8. Wu F, Zhao S, Yu B, Chen YM, Wang W, Song ZG, et al. A new coronavirus associated with human respiratory disease in China. *Nature* 2020;579(7798):265-9.
PUBMED | CROSSREF

9. Kim JM, Chung YS, Jo HJ, Lee NJ, Kim MS, Woo SH, et al. Identification of coronavirus isolated from a patient in Korea with COVID-19. *Osong Public Health Res Perspect* 2020;11(1):3-7.
PUBMED | CROSSREF
10. Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus-Infected Pneumonia. *N Engl J Med* 2020;382(13):1199-207.
PUBMED | CROSSREF
11. Pung R, Chiew CJ, Young BE, Chin S, Chen MI, Clapham HE, et al. Investigation of three clusters of COVID-19 in Singapore: implications for surveillance and response measures. *Lancet* 2020;395(10229):1039-46.
PUBMED | CROSSREF
12. Nishiura H, Linton NM, Akhmetzhanov AR. Serial interval of novel coronavirus (COVID-19) infections. *Int J Infect Dis* 2020;93:284-6.
PUBMED | CROSSREF
13. Ki M; Task Force for 2019-nCoV. Epidemiologic characteristics of early cases with 2019 novel coronavirus (2019-nCoV) disease in Korea. *Epidemiol Health* 2020;42:e2020007.
PUBMED | CROSSREF
14. Park YJ, Choe YJ, Park O, Park SY, Kim YM, Kim J, et al. Contact tracing during coronavirus disease outbreak, South Korea, 2020. *Emerg Infect Dis* 2020;26(10):2465-8.
PUBMED | CROSSREF
15. Park SY, Kim YM, Yi S, Lee S, Na BJ, Kim CB, et al. Coronavirus disease outbreak in call center, South Korea. *Emerg Infect Dis* 2020;26(8):1666-70.
PUBMED | CROSSREF
16. World Health Organization. Scientific brief. Modes of transmission of virus causing COVID-19: implications for IPC precaution recommendations. https://www.who.int/news-room/commentaries/detail/modes-of-transmission-of-virus-causing-covid-19-implications-for-ipc-precaution-recommendations. Updated 2020. Accessed July 12, 2020.
17. Wei J, Li Y. Airborne spread of infectious agents in the indoor environment. *Am J Infect Control* 2016;44(9 Suppl):S102-8.
PUBMED | CROSSREF
18. van Doremalen N, Bushmaker T, Morris DH, Holbrook MG, Gamble A, Williamson BN, et al. Aerosol and surface stability of SARS-CoV-2 as compared with SARS-CoV-1. *N Engl J Med* 2020;382(16):1564-7.
PUBMED | CROSSREF
19. Morawska L, Milton DK. It is time to address airborne transmission of COVID-19. *Clin Infect Dis*. Forthcoming 2020. DOI: 10.1093/cid/ciaa939.
PUBMED | CROSSREF
20. Dbouk T, Drikakis D. On coughing and airborne droplet transmission to humans. *Phys Fluids (1994)* 2020;32(5):053310.
PUBMED | CROSSREF
21. Lu J, Gu J, Li K, Xu C, Su W, Lai Z, et al. COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020. *Emerg Infect Dis* 2020;26(7):1628-31.
PUBMED | CROSSREF
22. European Center for Disease Prevention and Control. Heating, ventilation and air-conditioning systems in the context of COVID-19 (22 June 2020). https://www.ecdc.europa.eu/en/publications-data/heating-ventilation-air-conditioning-systems-covid-19. Updated 2020. Accessed November 5, 2020.

Exhibit E

Morbidity and Mortality Weekly Report

Community and Close Contact Exposures Associated with COVID-19 Among Symptomatic Adults ≥18 Years in 11 Outpatient Health Care Facilities — United States, July 2020

Kiva A. Fisher, PhD[1]; Mark W. Tenforde, MD, PhD[1,2]; Leora R. Feldstein, PhD[1]; Christopher J. Lindsell, PhD[3,4]; Nathan I. Shapiro, MD[3,5]; D. Clark Files, MD[3,6]; Kevin W. Gibbs, MD[3,6]; Heidi L. Erickson, MD[3,7]; Matthew E. Prekker, MD[3,7]; Jay S. Steingrub, MD[3,8]; Matthew C. Exline, MD[3,9]; Daniel J. Henning, MD[3,10]; Jennifer G. Wilson, MD[3,11]; Samuel M. Brown, MD[3,12]; Ithan D. Peltan, MD[3,12]; Todd W. Rice, MD[3,4]; David N. Hager, MD, PhD[3,13]; Adit A. Ginde, MD[3,14]; H. Keipp Talbot, MD[3,4]; Jonathan D. Casey, MD[3,4]; Carlos G. Grijalva, MD[3,4]; Brendan Flannery, PhD[1]; Manish M. Patel, MD[1]; Wesley H. Self, MD[3,4]; IVY Network Investigators; CDC COVID-19 Response Team

Community and close contact exposures continue to drive the coronavirus disease 2019 (COVID-19) pandemic. CDC and other public health authorities recommend community mitigation strategies to reduce transmission of SARS-CoV-2, the virus that causes COVID-19 (*1*,*2*). Characterization of community exposures can be difficult to assess when widespread transmission is occurring, especially from asymptomatic persons within inherently interconnected communities. Potential exposures, such as close contact with a person with confirmed COVID-19, have primarily been assessed among COVID-19 cases, without a non-COVID-19 comparison group (*3*,*4*). To assess community and close contact exposures associated with COVID-19, exposures reported by case-patients (154) were compared with exposures reported by control-participants (160). Case-patients were symptomatic adults (persons aged ≥18 years) with SARS-CoV-2 infection confirmed by reverse transcription–polymerase chain reaction (RT-PCR) testing. Control-participants were symptomatic outpatient adults from the same health care facilities who had negative SARS-CoV-2 test results. Close contact with a person with known COVID-19 was more commonly reported among case-patients (42%) than among control-participants (14%). Case-patients were more likely to have reported dining at a restaurant (any area designated by the restaurant, including indoor, patio, and outdoor seating) in the 2 weeks preceding illness onset than were control-participants (adjusted odds ratio [aOR] = 2.4; 95% confidence interval [CI] = 1.5–3.8). Restricting the analysis to participants without known close contact with a person with confirmed COVID-19, case-patients were more likely to report dining at a restaurant (aOR = 2.8, 95% CI = 1.9–4.3) or going to a bar/coffee shop (aOR = 3.9, 95% CI = 1.5–10.1) than were control-participants. Exposures and activities where mask use and social distancing are difficult to maintain, including going to places that offer on-site eating or drinking, might be important risk factors for acquiring COVID-19. As communities reopen, efforts to reduce possible exposures at locations that offer on-site eating and drinking options should be considered to protect customers, employees, and communities.

This investigation included adults aged ≥18 years who received a first test for SARS-CoV-2 infection at an outpatient testing or health care center at one of 11 Influenza Vaccine Effectiveness in the Critically Ill (IVY) Network sites* during July 1–29, 2020 (*5*). A COVID-19 case was confirmed by RT-PCR testing for SARS-CoV-2 RNA from respiratory specimens. Assays varied among facilities. Each site generated lists of adults tested within the study period by laboratory result; adults with laboratory-confirmed COVID-19 were selected by random sampling as case-patients. For each case-patient, two adults with negative SARS-CoV-2 RT-PCR test results were randomly selected as control-participants and matched by age, sex, and study location. After randomization and matching, 615 potential case-patients and 1,212 control-participants were identified and contacted 14–23 days after the date they received SARS-CoV-2 testing. Screening questions were asked to identify eligible adults. Eligible adults for the study were symptomatic at the time of their first SARS-CoV-2 test.

CDC personnel administered structured interviews in English or five other languages† by telephone and entered data into REDCap software (*6*). Among 802 adults contacted and who agreed to participate (295 case-patients and 507 control-participants), 332 reported symptoms at the time of initial SARS-CoV-2 testing and were enrolled in the study. Eighteen interviews were excluded because of nonresponse to the community exposure questions. The final analytic sample (314) included 154 case-patients (positive SARS-CoV-2 test results) and 160 control-participants (negative SARS-CoV-2

* Baystate Medical Center, Springfield, Massachusetts; Beth Israel Deaconess Medical Center, Boston, Massachusetts; University of Colorado School of Medicine, Aurora, Colorado; Hennepin County Medical Center, Minneapolis, Minnesota; Intermountain Healthcare, Salt Lake City, Utah; Ohio State University Wexner Medical Center, Columbus, Ohio; Wake Forest University Baptist Medical Center, Winston-Salem, North Carolina; Vanderbilt University Medical Center, Nashville, Tennessee; John Hopkins Hospital, Baltimore, Maryland; Stanford University Medical Center, Palo Alto, California; University of Washington Medical Center, Seattle, Washington). Participating states include California, Colorado, Maryland, Massachusetts, Minnesota, North Carolina, Ohio, Tennessee, Utah, and Washington.

† Other languages included Spanish, Arabic, Vietnamese, Portuguese, and Russian.

test results). Among nonparticipants, 470 were ineligible (i.e., were not symptomatic or had multiple tests), and 163 refused to participate. This activity was reviewed by CDC and participating sites and conducted consistent with applicable federal law and CDC policy.[§]

Data collected included demographic characteristics, information on underlying chronic medical conditions,[¶] symptoms, convalescence (self-rated physical and mental health), close contact (within 6 feet for ≥15 minutes) with a person with known COVID-19, workplace exposures, mask-wearing behavior, and community activities ≤14 days before symptom onset. Participants were asked about wearing a mask and possible community exposure activities (e.g., gatherings with ≤10 or >10 persons in a home; shopping; dining at a restaurant; going to an office setting, salon, gym, bar/coffee shop, or church/religious gathering; or using public transportation) on a five-point Likert-type scale ranging from "never" to "more than once per day" or "always"; for analysis, community activity responses were dichotomized as never versus one or more times during the 14 days before illness onset. For each reported activity, participants were asked to quantify degree of adherence to recommendations such as wearing a face mask of any kind or social distancing among other persons at that location, with response options ranging from "none" to "almost all." Descriptive and statistical analyses were performed to compare case-patients with control-participants, assessing differences in demographic characteristics, community exposures, and close contact. Although an effort was made initially to match case-patients to control-participants based on a 1:2 ratio, not all potential participants were eligible or completed an interview, and therefore an unmatched analysis was performed. Unconditional logistic regression models with generalized estimating equations with exchangeable correlation structure correcting standard error estimates for site-level clustering were used to assess differences in community exposures between case-patients and control-participants, adjusting for age, sex, race/ethnicity, and presence of one or more underlying chronic medical conditions. In each model, SARS-CoV-2 test result (i.e., positive or negative) was the outcome variable, and each community exposure activity was the predictor variable. The first model included the full analytic sample (314). A second model was restricted to participants who did not report close contact to a person with COVID-19 (89 case-patients and 136 control-participants). Statistical analyses were conducted using SAS software (version 9.4; SAS Institute).

[§] Activity was determined to meet the requirements of public health surveillance as defined in 45 CFR 46.102(l)(2).

[¶] Cardiac condition, hypertension, asthma, chronic obstructive pulmonary disease, immunodeficiency, psychiatric condition, diabetes, or obesity.

Compared with case-patients, control-participants were more likely to be non-Hispanic White ($p<0.01$), have a college degree or higher ($p<0.01$), and report at least one underlying chronic medical condition ($p = 0.01$) (Table). In the 14 days before illness onset, 71% of case-patients and 74% of control-participants reported always using cloth face coverings or other mask types when in public. Close contact with one or more persons with known COVID-19 was reported by 42% of case-patients compared with 14% of control-participants ($p<0.01$), and most (51%) close contacts were family members.

Approximately one half of all participants reported shopping and visiting others inside a home (in groups of ≤10 persons) on ≥1 day during the 14 days preceding symptom onset. No significant differences were observed in the bivariate analysis between case-patients and control-participants in shopping; gatherings with ≤10 persons in a home; going to an office setting; going to a salon; gatherings with >10 persons in a home; going to a gym; using public transportation; going to a bar/coffee shop; or attending church/religious gathering. However, case-patients were more likely to have reported dining at a restaurant (aOR = 2.4, 95% CI = 1.5–3.8) in the 2 weeks before illness onset than were control-participants (Figure). Further, when the analysis was restricted to the 225 participants who did not report recent close contact with a person with known COVID-19, case-patients were more likely than were control-participants to have reported dining at a restaurant (aOR = 2.8, 95% CI = 1.9–4.3) or going to a bar/coffee shop (aOR = 3.9, 95% CI = 1.5–10.1). Among 107 participants who reported dining at a restaurant and 21 participants who reported going to a bar/coffee shop, case-patients were less likely to report observing almost all patrons at the restaurant adhering to recommendations such as wearing a mask or social distancing ($p = 0.03$ and $p = 0.01$, respectively).

Discussion

In this investigation, participants with and without COVID-19 reported generally similar community exposures, with the exception of going to locations with on-site eating and drinking options. Adults with confirmed COVID-19 (case-patients) were approximately twice as likely as were control-participants to have reported dining at a restaurant in the 14 days before becoming ill. In addition to dining at a restaurant, case-patients were more likely to report going to a bar/coffee shop, but only when the analysis was restricted to participants without close contact with persons with known COVID-19 before illness onset. Reports of exposures in restaurants have been linked to air circulation (*7*). Direction, ventilation, and intensity of airflow might affect virus transmission, even if social distancing measures and mask use are implemented according to current guidance. Masks cannot

TABLE. Characteristics of symptomatic adults ≥18 years who were outpatients in 11 academic health care facilities and who received positive and negative SARS-CoV-2 test results (N = 314)* — United States, July 1–29, 2020

Characteristic	No. (%) Case-patients (n = 154)	Control participants (n = 160)	P-value
Age group, yrs			
18–29	44 (28.6)	39 (24.4)	0.18
30–44	46 (29.9)	62 (38.7)	
45–59	46 (29.9)	35 (21.9)	
≥60	18 (11.7)	24 (15.0)	
Sex			
Men	75 (48.7)	72 (45.0)	0.51
Women	79 (51.3)	88 (55.0)	
Race/Ethnicity†			
White, non-Hispanic	92 (59.7)	124 (77.5)	<0.01
Hispanic/Latino	29 (18.8)	12 (7.5)	
Black, non-Hispanic	27 (17.5)	19 (11.9)	
Other, non-Hispanic	6 (3.9)	5 (3.1)	
Education (missing = 3)			
Less than high school	16 (10.5)	3 (1.9)	<0.01
High school degree or some college	60 (39.2)	48 (30.4)	
College degree or more	77 (50.3)	107 (67.7)	
At least one underlying chronic medical condition§	**75 (48.7)**	**98 (61.2)**	**0.01**
Community exposure 14 days before illness onset¶			
Shopping	131 (85.6)	141 (88.1)	0.51
Home, ≤10 persons	79 (51.3)	84 (52.5)	0.83
Restaurant	63 (40.9)	44 (27.7)	0.01
Office setting	37 (24.0)	47 (29.6)	0.27
Salon	24 (15.6)	28 (17.6)	0.63
Home, >10 persons	21 (13.6)	24 (15.0)	0.73
Gym	12 (7.8)	10 (6.3)	0.60
Public transportation	8 (5.2)	10 (6.3)	0.68
Bar/Coffee shop	13 (8.5)	8 (5.0)	0.22
Church/Religious gathering	12 (7.8)	8 (5.0)	0.32
Restaurant: others following recommendations such as wearing a face covering or mask of any kind or social distancing (n = 107)			
None/A few	12 (19.0)	1 (2.3)	0.03
About half/Most	25 (39.7)	21 (47.7)	
Almost all	26 (41.3)	22 (50.0)	
Bar: others following recommendations such as wearing a face covering or mask of any kind or social distancing (n = 21)			
None/A few	4 (31.8)	2 (25.0)	0.01
About half/Most	7 (53.8)	0 (0.0)	
Almost all	2 (15.4)	6 (75.0)	

See table footnotes on the next page.

be effectively worn while eating and drinking, whereas shopping and numerous other indoor activities do not preclude mask use.

Among adults with COVID-19, 42% reported close contact with a person with COVID-19, similar to what has been reported previously (*4*). Most close contact exposures were to family members, consistent with household transmission of SARS-CoV-2 (*8*). Fewer (14%) persons who received a negative SARS-CoV-2 test result reported close contact with a person with known COVID-19. To help slow the spread of SARS-CoV-2, precautions should be implemented to stay home once exposed to someone with COVID-19,** in addition to adhering to recommendations to wash hands often, wear masks, and social distance.†† If a family member or other close contact is ill, additional prevention measures can be taken to reduce transmission, such as cleaning and disinfecting the home, reducing shared meals and items, wearing gloves, and wearing masks, for those with and without known COVID-19.§§

The findings in this report are subject to at least five limitations. First, the sample included 314 symptomatic patients who actively sought testing during July 1–29, 2020 at 11 health care facilities. Symptomatic adults with negative SARS-CoV-2 test results might have been infected with other respiratory

** https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html.

†† https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/index.html.

§§ https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/index.html.

TABLE. (*Continued*) Characteristics of symptomatic adults ≥18 years who were outpatients in 11 academic health care facilities and who received positive and negative SARS-CoV-2 test results (N = 314)* — United States, July 1–29, 2020

Characteristic	No. (%) Case-patients (n = 154)	Control participants (n = 160)	P-value
Previous close contact with a person with known COVID-19 (missing = 1)			
No	89 (57.8)	136 (85.5)	<0.01
Yes	65 (42.2)	23 (14.5)	
Relationship to close contact with known COVID-19 (n = 88)			
Family	33 (50.8)	5 (21.7)	<0.01
Friend	9 (13.8)	4 (17.4)	
Work colleague	11 (16.9)	6 (26.1)	
Other**	6 (9.2)	8 (34.8)	
Multiple	6 (9.2)	0 (0.0)	
Reported use of cloth face covering or mask 14 days before illness onset (missing = 2)			
Never	6 (3.9)	5 (3.1)	0.86
Rarely	6 (3.9)	6 (3.8)	
Sometimes	11 (7.2)	7 (4.4)	
Often	22 (14.4)	23 (14.5)	
Always	108 (70.6)	118 (74.2)	

* Respondents who completed the interview 14–23 days after their test date. Five participants had significant missingness for exposure questions and were removed from the analysis. Patients were randomly sampled from 11 academic health care systems that are part of the Influenza Vaccine Effectiveness in the Critically Ill Network sites (Baystate Medical Center, Springfield, Massachusetts; Beth Israel Deaconess Medical Center, Boston, Massachusetts; University of Colorado School of Medicine, Aurora, Colorado; Hennepin County Medical Center, Minneapolis, Minnesota; Intermountain Healthcare, Salt Lake City, Utah; Ohio State University Wexner Medical Center, Columbus, Ohio; Wake Forest University Baptist Medical Center, Winston-Salem, North Carolina; Vanderbilt University Medical Center, Nashville, Tennessee; John Hopkins Hospital, Baltimore, Maryland; Stanford University Medical Center, Palo Alto, California; University of Washington Medical Center, Seattle, Washington). Participating states include California, Colorado, Maryland, Massachusetts, Minnesota, North Carolina, Ohio, Tennessee, Utah, and Washington.

† Other race includes responses of Native American/Alaska Native, Asian, Native Hawaiian/Other Pacific Islander, and other; these were combined because of small sample sizes.

§ Reported at least one of the following underlying chronic medical conditions: cardiac condition, hypertension, asthma, chronic obstructive pulmonary disease, immunodeficiency, psychiatric condition, diabetes, or obesity.

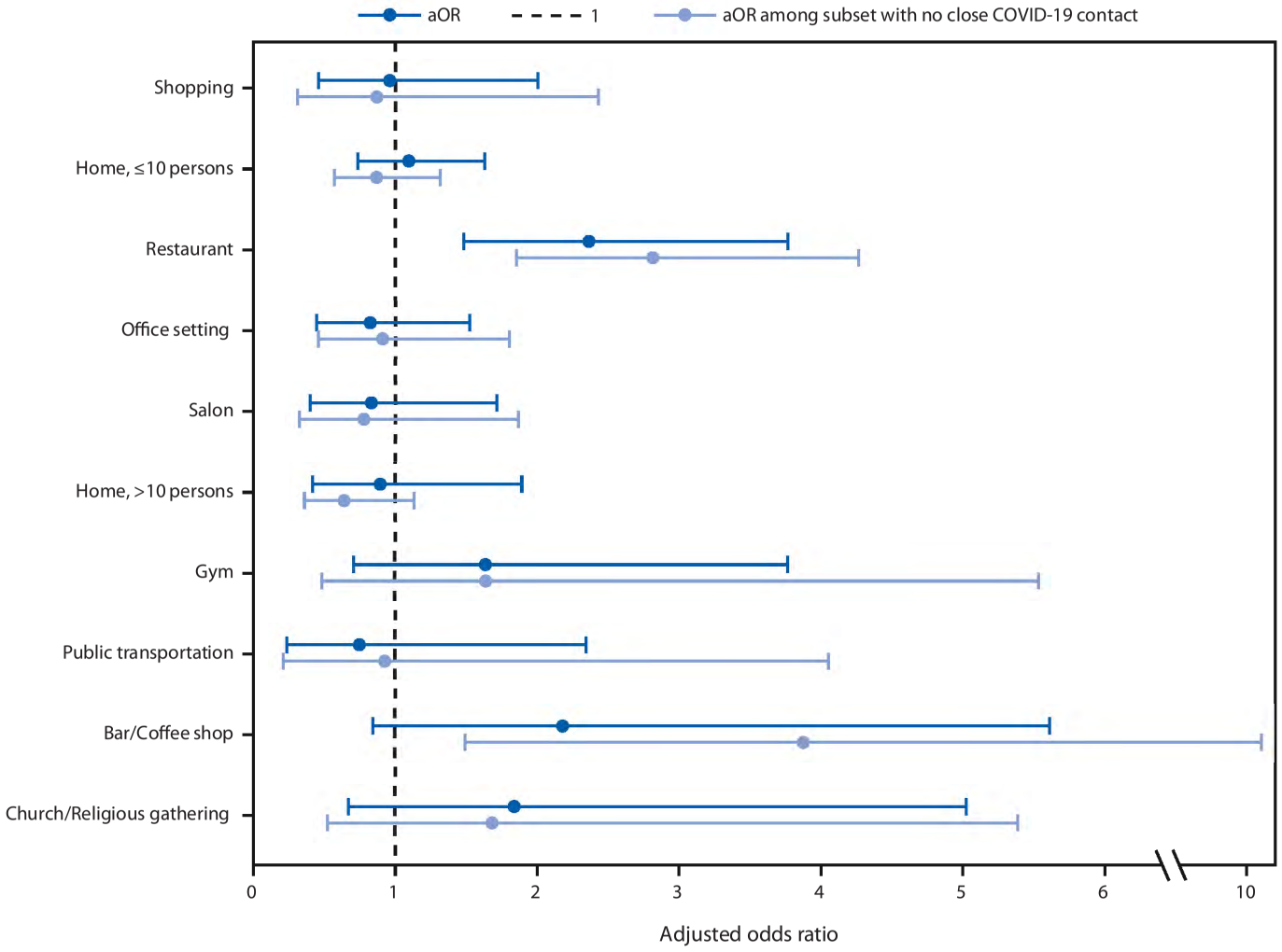
¶ Community exposure questions asked were "In the 14 days before feeling ill about how often did you:" with options of "shop for items (groceries, prescriptions, home goods, clothing, etc.)" (missing = 1); "have people visit you inside your home or go inside someone else's home where there were more than 10 people"; "have people visit you inside your home or go inside someone else's home where there were 10 people or less"; "go to church or a religious gathering/place of worship" (missing = 1); "go to a restaurant (dine-in, any area designated by the restaurant including patio seating)" (missing = 1); "go to a bar or coffee shop (indoors)" (missing = 2); "use public transportation (bus, subway, streetcar, train, etc.)" (missing = 1); "go to an office setting (other than for healthcare purposes)" (missing = 1); "go to a gym or fitness center" (missing = 1); and "go to a salon or barber (e.g., hair salon, nail salon, etc.)" (missing = 1). Response options were coded as never versus at least once in the 14 days prior to illness onset. Some participants had missing data for exposure questions:

** Other includes patients of health care workers (9), patron of a restaurant (1), spouse of employee (1), day care teacher (1), member of a religious congregation (1), and unspecified (1).

viruses and had similar exposures to persons with cases of such illnesses. Persons who did not respond, or refused to participate, could be systematically different from those who were interviewed for this investigation. Efforts to age- and sex-match participating case-patients and control-participants were not maintained because of participants not meeting the eligibility criteria, refusing to participate, or not responding, and this was accounted for in the analytic approach. Second, unmeasured confounding is possible, such that reported behaviors might represent factors, including concurrently participating in activities where possible exposures could have taken place, that were not included in the analysis or measured in the survey. Of note, the question assessing dining at a restaurant did not distinguish between indoor and outdoor options. In addition, the question about going to a bar or coffee shop did not distinguish between the venues or service delivery methods, which might represent different exposures. Third, adults in the study were from one of 11 participating health care facilities and might not be representative of the United States population. Fourth, participants were aware of their SARS-CoV-2 test results, which could have influenced their responses to questions about community exposures and close contacts. Finally, case or control status might be subject to misclassification because of imperfect sensitivity or specificity of PCR-based testing (*9*,*10*).

This investigation highlights differences in community and close contact exposures between adults who received a positive SARS-CoV-2 test result and those who received a negative SARS-CoV-2 test result. Continued assessment of various types of activities and exposures as communities, schools, and workplaces reopen is important. Exposures and activities where mask use and social distancing are difficult to maintain, including going to locations that offer on-site eating and drinking, might be important risk factors for

FIGURE. Adjusted odds ratio (aOR)* and 95% confidence intervals for community exposures† associated with confirmed COVID-19 among symptomatic adults aged ≥18 years (N = 314) — United States, July 1–29, 2020



Abbreviation: COVID-19 = coronavirus disease 2019.

* Adjusted for race/ethnicity, sex, age, and reporting at least one underlying chronic medical condition. Odds ratios were estimated using unconditional logistic regression with generalized estimating equations, which accounted for Influenza Vaccine Effectiveness in the Critically Ill Network site-level clustering. A second model was restricted to participants who did not report close contact to a person known to have COVID-19 (n = 225).

† Community exposure questions asked were "In the 14 days before feeling ill about how often did you: shop for items (groceries, prescriptions, home goods, clothing, etc.); have people visit you inside your home or go inside someone else's home where there were more than 10 people; have people visit you inside your home or go inside someone else's home where there were 10 people or less; go to church or a religious gathering/place of worship; go to a restaurant (dine-in, any area designated by the restaurant including patio seating); go to a bar or coffee shop (indoors); use public transportation (bus, subway, streetcar, train, etc.); go to an office setting (other than for healthcare purposes); go to a gym or fitness center; go to a salon or barber (e.g., hair salon, nail salon, etc.)." Response options were coded as never versus at least once in the 14 days before illness onset.

SARS-CoV-2 infection. Implementing safe practices to reduce exposures to SARS-CoV-2 during on-site eating and drinking should be considered to protect customers, employees, and communities[¶¶] and slow the spread of COVID-19.

[¶¶] https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/personal-social-activities.html#restaurant; https://www.cdc.gov/coronavirus/2019-ncov/community/organizations/business-employers/bars-restaurants.html; https://www.cdc.gov/coronavirus/2019-ncov/images/community/Rest_Bars_RiskAssessment.jpg.

Acknowledgments

Zhanar Haimovich, Northrop Grumman; Sherri Pals, Division of Global HIV & TB, Center for Global Health, CDC.

Corresponding author: Kiva A. Fisher, eocevent458@cdc.gov.

Summary

What is already known about the topic?

Community and close contact exposures contribute to the spread of COVID-19.

What is added by this report?

Findings from a case-control investigation of symptomatic outpatients from 11 U.S. health care facilities found that close contact with persons with known COVID-19 or going to locations that offer on-site eating and drinking options were associated with COVID-19 positivity. Adults with positive SARS-CoV-2 test results were approximately twice as likely to have reported dining at a restaurant than were those with negative SARS-CoV-2 test results.

What are the implications for public health practice?

Eating and drinking on-site at locations that offer such options might be important risk factors associated with SARS-CoV-2 infection. Efforts to reduce possible exposures where mask use and social distancing are difficult to maintain, such as when eating and drinking, should be considered to protect customers, employees, and communities.

[1]CDC COVID-19 Response Team; [2]Epidemic Intelligence Service, CDC; [3]Influenza Vaccine Effectiveness in the Critically Ill (IVY) Network; [4]Vanderbilt University Medical Center, Nashville, Tennessee; [5]Beth Israel Deaconess Medical Center, Boston, Massachusetts; [6]Wake Forest University Baptist Medical Center, Winston-Salem, North Carolina; [7]Hennepin County Medical Center, Minneapolis, Minnesota; [8]Baystate Medical Center, Springfield, Massachusetts; [9]Ohio State University Wexner Medical Center, Columbus, Ohio; [10]University of Washington Medical Center, Seattle, Washington; [11]Stanford University Medical Center, Palo Alto, California; [12]Intermountain Healthcare, Salt Lake City, Utah; [13]Johns Hopkins Hospital, Baltimore, Maryland; [14]University of Colorado School of Medicine, Aurora, Colorado.

IVY Network Investigators

Kimberly W. Hart, Vanderbilt University Medical Center; Robert McClellan, Vanderbilt University Medical Center; Hsi-nien Tan, Vanderbilt University Medical Center; Adrienne Baughman, Vanderbilt University Medical Center.

CDC COVID-19 Response Team

Nora A. Hennesy, CDC COVID-19 Response Team; Brittany Grear, CDC COVID-19 Response Team; Michael Wu, CDC COVID-19 Response Team; Kristin Mlynarczyk, CDC COVID-19 Response Team; Luc Marzano, CDC COVID-19 Response Team; Zuwena Plata, CDC COVID-19 Response Team; Alexis Caplan, CDC COVID-19 Response Team; Samantha M. Olson, CDC COVID-19 Response Team; Constance E. Ogokeh, CDC COVID-19 Response Team; Emily R. Smith, CDC COVID-19 Response Team; Sara S. Kim, CDC COVID-19 Response Team; Eric P. Griggs, CDC COVID-19 Response Team; Bridget Richards, CDC COVID-19 Response Team; Sonya Robinson, CDC COVID-19 Response Team; Kaylee Kim, CDC COVID-19 Response Team; Ahmed M. Kassem, CDC COVID-19 Response Team; Courtney N. Sciarratta, CDC COVID-19 Response Team; Paula L. Marcet, CDC COVID-19 Response Team.

All authors have completed and submitted the International Committee of Medical Journal Editors form for disclosure of potential conflicts of interest. Carlos G. Grijalva reports grants from Campbell Alliance, the National Institutes of Health, the Food and Drug Administration, the Agency for Health Care Research and Quality and Sanofi-Pasteur, and consultation fees from Pfizer, Merck, and Sanofi-Pasteur. Christopher J. Lindsell reports grants from National Institutes of Health and the Department of Defense and other support from Marcus Foundation, Endpoint Health, Entegrion, bioMerieux, and Bioscape Digital, outside the submitted work. Nathan I. Shapiro reports grants from the National Institutes of Health, Rapid Pathogen Screening, Inflammatix, and Baxter, outside the submitted work. Daniel J. Henning reports personal fees from CytoVale and grants from Baxter, outside the submitted work. Samuel M. Brown reports grants from National Institutes of Health, Department of Defense, Intermountain Research and Medical Foundation, and Janssen and consulting fees paid to his employer from Faron and Sedana, outside the submitted work. Ithan D. Peltan reports grants from the National Institutes of Health, Asahi Kasei Pharma, Immunexpress Inc., Janssen Pharmaceuticals, and Regeneron, outside the submitted work. Todd W. Rice reports personal fees from Cumberland Pharmaceuticals, Inc, Cytovale, Inc, and Avisa, LLC, outside the submitted work. Adit A. Ginde reports grants from the National Institutes of Health and Department of Defense, outside the submitted work. H. Keipp Talbot reports serving on the Data Safety Monitoring Board for Seqirus. No other potential conflicts of interest were disclosed.

References

1. CDC. Coronavirus disease 2019 (COVID-19): implementation of mitigation strategies for communities with local COVID-19 transmission. Atlanta, GA: US Department of Health and Human Services; 2020. https://www.cdc.gov/coronavirus/2019-ncov/community/community-mitigation.html
2. CDC. Coronavirus disease 2019 (COVID-19): community, work, and school: information for where you live, work, learn, and play. Atlanta, GA: US Department of Health and Human Services, CDC; 2020. https://www.cdc.gov/coronavirus/2019-ncov/community/index.html
3. Marshall K, Vahey GM, McDonald E, et al.; Colorado Investigation Team. Exposures before issuance of stay-at-home orders among persons with laboratory-confirmed COVID-19—Colorado, March 2020. MMWR Morb Mortal Wkly Rep 2020;69:847–9. https://doi.org/10.15585/mmwr.mm6926e4
4. Tenforde MW, Billig Rose E, Lindsell CJ, et al.; CDC COVID-19 Response Team. Characteristics of adult outpatients and inpatients with COVID-19—11 academic medical centers, United States, March–May 2020. MMWR Morb Mortal Wkly Rep 2020;69:841–6. https://doi.org/10.15585/mmwr.mm6926e3
5. Stubblefield WB, Talbot HK, Feldstein L, et al.; Influenza Vaccine Effectiveness in the Critically Ill (IVY) Investigators. Seroprevalence of SARS-CoV-2 among frontline healthcare personnel during the first month of caring for COVID-19 patients—Nashville, Tennessee. Clin Infect Dis 2020;ciaa936. https://doi.org/10.1093/cid/ciaa936
6. Harris PA, Taylor R, Minor BL, et al.; REDCap Consortium. The REDCap consortium: building an international community of software platform partners. J Biomed Inform 2019;95:103208. https://doi.org/10.1016/j.jbi.2019.103208

7. Lu J, Gu J, Li K, et al. COVID-19 outbreak associated with air conditioning in restaurant, Guangzhou, China, 2020. Emerg Infect Dis 2020;26:1628–31. https://doi.org/10.3201/eid2607.200764
8. Lei H, Xu X, Xiao S, Wu X, Shu Y. Household transmission of COVID-19-a systematic review and meta-analysis. J Infect 2020. Epub August 25, 2020. https://doi.org/10.1016/j.jinf.2020.08.033
9. Sethuraman N, Jeremiah SS, Ryo A. Interpreting diagnostic tests for SARS-CoV-2. JAMA 2020;323:2249–51. https://doi.org/10.1001/jama.2020.8259
10. Tahamtan A, Ardebili A. Real-time RT-PCR in COVID-19 detection: issues affecting the results. Expert Rev Mol Diagn 2020;20:453–4. https://doi.org/10.1080/14737159.2020.1757437

Erratum

Vol. 69, No. 36

In the report "Community and Close Contact Exposures Associated with COVID-19 Among Symptomatic Adults ≥18 Years in 11 Outpatient Health Care Facilities — United States, July 2020," on page 1262, the e-mail for contact information has been updated to **eocevent101@cdc.gov**.