Title

Delay in the Effect of Restricting Community Mobility on the Spread of COVID-19 During the First Wave in the United States.

Journal

Open Forum Infectious Diseases, 9(1)

ISSN

2328-8957

Authors

He, Shan
Lee, Jooyoung
Langworthy, Benjamin
et al.

Publication Date

2022

DOI

10.1093/ofid/ofab586

Peer reviewed

Open Forum Infectious Diseases

MAJOR ARTICLE

Delay in the Effect of Restricting Community Mobility on the Spread of COVID-19 During the First Wave in the United States

Shan He,[1,2,a] Jooyoung Lee,[3,a] Benjamin Langworthy,[4] Junyi Xin,[4] Peter James,[5,6] Yang Yang,[7,b] and Molin Wang[1,4,8,b]

[1]Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA, [2]Division of Sleep and Circadian Disorders, Brigham and Women's Hospital, Boston, Massachusetts, USA, [3]Department of Applied Statistics, Chung-Ang University, Seoul, Korea, [4]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA, [5]Department of Environmental Health, Harvard T.H. Chan School of Public Health, Boston, Massachusetts, USA, [6]Department of Population Medicine, Harvard Medical School and Harvard Pilgrim Health Care Institute, Boston, Massachusetts, USA, [7]Department of Biostatistics, University of Florida, College of Public Health and Health Professions & Emerging Pathogens Institute, Gainesville, Florida, USA, and [8]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, Massachusetts, USA

Background. It remains unclear how changes in human mobility shaped the transmission dynamic of coronavirus disease 2019 (COVID-19) during its first wave in the United States.

Methods. By coupling a Bayesian hierarchical spatiotemporal model with reported case data and Google mobility data at the county level, we found that changes in movement were associated with notable changes in reported COVID-19 incidence rates about 5 to 7 weeks later.

Results. Among all movement types, residential stay was the most influential driver of COVID-19 incidence rate, with a 10% increase 7 weeks ago reducing the disease incidence rate by 13% (95% credible interval, 6%–20%). A 10% increase in movement from home to workplaces, retail and recreation stores, public transit, grocery stores, and pharmacies 7 weeks ago was associated with an increase of 5%–8% in the COVID-10 incidence rate. In contrast, parks-related movement showed minimal impact.

Conclusions. Policy-makers should anticipate such a delay when planning intervention strategies restricting human movement.

Keywords. community mobility; COVID-19; infectious diseases; spatio-temporal models; statistical modeling.

In response to the coronavirus disease 2019 (COVID-19) pandemic, scientists have been trying to understand the roles of human activities, nonpharmaceutical interventions (NPIs), and socioenvironmental conditions in the transmission of COVID-19 [1–14], in parallel with the race toward effective vaccines. As pointed out in recent publications, NPIs such as closing citywide transportation and entertainment venues and banning public gatherings were associated with a marked reduction in the number of COVID-19 cases in China [1], and wearing facial masks was associated with reduced risk of COVID-19 [4]. Human mobility is a key driver of the spread of many respiratory pathogens including severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), and restriction of human mobility, used alone or in combination with other NPI options, was highly effective in slowing or controlling the spread of COVID-19 [1–3]. The effect of human movement changes on disease transmission dynamics, however, may take some time to become apparent [7]. In addition, the effect likely exhibits a certain level of spatial heterogeneity as human movement patterns heavily depend on environmental, cultural, and socioeconomic factors. In the United States, where COVID-19 hit the hardest during 2020, "shelter in place" and "stay at home" orders were implemented at different levels across states but were lifted within 2 months in most states (https://www.kaggle.com/lin0li/us-lockdown-dates-dataset), reflecting lack of evidence-based guidance for making these policies. Accurate quantification of the relationship between human movement and COVID-19 transmission at a fine geographic scale will help bridge the gap between scientific evidence and policy-making.

Although a large amount of research has focused on the impact of NPIs on the transmission dynamic of COVID-19, researchers have mostly focused on policies that affect human movement, such as travel restriction [3] and school closure [10]. There are few publications focusing directly on the impact of the level of human movement itself. Chang et al. modeled mobility networks and transmission of COVID-19 in 10 large US metropolitan areas and found that both magnitude and timing of mobility reduction were influential in the reduction of infections [8]. This research is limited to urban settings and

Received 17 June 2021; editorial decision 15 November 2021; accepted 18 November 2021; published online 22 November 2021.

[a]Equal contribution

[b]Equal contribution

Correspondence: Molin Wang, PhD, Department of Epidemiology and Biostatistics, Harvard T.H. Chan School of Public Health, Kresge Building Room 828, 677 Huntington Ave, Boston, MA 02115 (stmow@channing.harvard.edu).

Open Forum Infectious Diseases®2021

https://doi.org/10.1093/ofid/ofab586

imposed strong assumptions on under-reporting and delay of reporting to model the underlying transmission process. Using generalized linear models and Pearson correlation, Badr et al. [7] reported that mobility patterns are correlated with COVID-19 growth rates, although the impact may not be perceptible for up to 3 weeks. However, the study focused on only the 25 counties reporting the highest number of cases in the United States. Similarly, using integrated mobile data, Xiong et al. [9] reported a strong positive relationship between mobility inflow and the number of COVID-19 cases in the United States through log-linear regressions with lagged mobility covariates. However, none of these studies offer clear answers to the following questions: (1) How fast did the incidence rates change due to changes in human mobility of different types? And (2) did the impact of human mobility on disease incidence vary across the United States? In addition, potential overdispersion and spatial correlation were not fully addressed in the previous analyses. Ignoring these issues could bias our understanding of the role of human mobility and subsequently misguide intervention policies.

In this paper, we use Bayesian hierarchical spatiotemporal models to assess the potential temporal lag in the effect of particular human movement indices and reported incidence of COVID-19. We are also able to investigate whether the lag varies spatially. These models explicitly consider overdispersion in reported incidence and potential spatial autocorrelations among counties. The models are adjusted for potential risk factors of COVID-19 [15, 16], including temperature, age profiles, percentage of African Americans, percentage of Latinos, percentage below the poverty line, percent obese, percentage with Bachelor's degree or above, and population density.

METHODS

Data Sources and Structure

Supplementary Table 1 summarizes the sources of our data. Daily human movement data for 2949 counties in the United States from February 15, 2020, to July 15, 2020, were downloaded from Google. As of April 2018, Google Maps had 154.4 million users in the United States [17]. We consider community mobility indices corresponding to 6 categories: workplaces, residential complexes, retail and recreation stores, groceries and pharmacies, transit stations, and parks. In the COVID-19 Community Mobility Reports, residential movement is defined as the relative change in duration spent in residence, while other movement indices are defined as relative changes in daily number of visitors. The baseline values for calculating all relative changes are the median values during the 5-week period from January 3 to February 6, 2020 [18]. All categories except residential encompass a wide range of destinations. For example, park destinations include local/national parks, public beaches, dog parks, and plazas. Transit destinations include subways, buses, and trains. Retail/recreation destinations include restaurants, cafés, shopping malls, amusement parks, museums, and cinemas. The mobility data for trips related to work, retail and recreation, and grocery and pharmacy are available for 84.4%–94.8% of the US counties, but fewer data are available for trips related to transit (34.2%), parks (39.4%), or time spent in residence (53.3%). Each mobility index was analyzed individually, excluding counties with missing values. For robust statistical inference, we further excluded counties that reported <50 cumulative cases of COVID-19 as of July 15, 2020, or that reported biweekly rather than weekly incidence.

County-level weekly reported numbers of COVID-19 cases in the United States were extracted from the "COVID-19 Data Repository by the Center for Systems Science and Engineering" (CSSE) at Johns Hopkins University [19]. According to guidance from the Centers for Disease Control and Prevention (CDC) published on April 14, 2020 [20], county-reported numbers of COVID-19 cases include both laboratory-confirmed cases and probable cases. A probable case is defined as meeting 1 of the following: (1) clinical criteria and epidemiologic evidence; (2) presumptive laboratory evidence and either clinical criteria or epidemiologic evidence; (3) vital records criteria. A lab-confirmed case is defined as a case with confirmatory laboratory evidence. The end of our study period was July 15, 2020, for all counties, when hospitals stopped directly reporting COVID-19 data to the CDC [21].

We collected county-level time-independent socioenvironmental variables that were potentially associated with COVID-19 incidence from multiple data sources (Supplementary Table 1). The following variables were extracted from the American Community Survey: population density (measured in quantiles among all US counties), proportion of population in 3 age groups (20–45, 45–65, and >65 years), percentage below the poverty line, percentage of African American residents, percentage of Hispanic residents, and percentage of population with a Bachelor's degree or above. Average winter and summer maximum temperatures in degrees Fahrenheit were obtained from Wu et al. [22]. County-level proportions of obese residents were obtained from the County Health Rankings & Roadmaps program of the University of Wisconsin Population Health Institute.

Statistical Model

To quantify the impact of human movement on COVID-19 incidence, we combined a Bayesian spatiotemporal generalized additive mixed model (GAMM) with a distributed lag model (DLM) [23]. The GAMM allows the effects of continuous risk factors to have flexible shapes, and the DLM permits lagged effects of human movement at each time point on future incidences. In addition, to account for spatial dependency between adjacent counties and the nonlinear time trends, spatial random effects and temporal nonlinear effects are included.

We estimated the association of weekly incidence of COVID-19 with weekly averages of mobility indices, while adjusting for potential confounders, spatiotemporal correlation, and spatiotemporal heterogeneity. Let Y_{it} be the reported number of COVID-19 cases and M_{it} be the average value of a mobility index during week t in county i. Let $j(i)$ indicate the state of county i, which is sometimes written as j when it is used to index states. Let Z_i be the time-independent covariates for county i, which includes all the county-level potential confounders mentioned above. The distributed lag model is structured as

$$Y_{it} \sim \text{Negative Binomial}\,(\mu_{it},\ \varphi)$$

$$\log(\mu_{it}) = \log(of\ fset_i) + \alpha_0 + Z_i^T\beta + \sum_{u=t-8}^{t} M_{iu}\gamma(t-u) + \sum_{u=t-8}^{t} M_{iu}\eta_{j(i)}(t-u) + f_S(i) + f_T(t) + f_{ST}(i,t), \quad (1)$$

where $E(Y_{it}) = \mu_{it}$ and $Var(Y_{it}) = \mu_{it} + \varphi\mu_{it}^2$ are the mean and variance of the weekly incidence Y_{it}, and $of\ fset_i$ is the county's population size. The intercept α_0 is assigned a Gaussian prior with mean 0 and precision 0, and the regression coefficients β are assigned Gaussian priors with mean 0 and precision 0.001 by the R-INLA default setting [24]. Assuming a maximum lag of 8 weeks, we consider 2 levels of distributed lag effects of the mobility index. The fixed effects at the national level are captured by $\sum_{u=t-8}^{t} M_{iu}\gamma(t-u)$, and the random effects at the state level are modeled as $\sum_{u=t-8}^{t} M_{iu}\eta_{j(i)}(t-u)$. The smooth effect curves $\gamma(t-u)$ and $\eta_{j(i)}(t-u)$, $j(i) = 1, \ldots, 46$, are each a linear combination of quadratic B-splines with 2 boundary knots, where 46is the number of states included. The coefficients of the B-splines are γ_k, $k = 1, 2, 3$ for the national level and η_{jk}, $j = 1, \ldots, 46$, $k = 1, 2, 3$, for the state level. We assign the following Gaussian priors to the coefficients for the B-splines: $\gamma_k \sim Normal(0, 0.001)$, $\eta_{jk} \sim Normal(0, \tau_{\eta j}^{-1})$, where $\tau_{\eta j} \sim Gamma(1, 0.001)$ for $j = 1, \ldots, 46$. When a single lag of u weeks rather than the summation of lags is considered, the summations of lag effects in the model are replaced by $\gamma M_{i,t-u} + \eta_{j(i)} M_{i,t-u}$, where γ and $\eta_{j(i)}$ are scalar coefficients. We report incidence rate ratios (IRRs) per 10% increase in each movement index. For models with a single lag, $\exp(\gamma)$ is the national IRR, and $\exp(\gamma + \eta_j)$ is the IRR for state j. For models with distributed lags, the IRR curve as a function of lag δ is calculated as $\exp(\gamma(\delta))$ for the national average and $\exp(\gamma(\delta) + \eta_j(\delta))$ for state j.

The spatial effect $f_S(i)$ is modeled as $f_S(i) = u_i + v_i$. The spatially unstructured random effect u_i is assigned a Gaussian prior, with mean 0 and precision $\tau_u \sim Gamma(1, 0.001)$. The spatially structured effect v_i was modeled by an intrinsic conditional autoregressive prior (ICAR) [25] to account for spatial autocorrelation, $v_i | v_{i'} \sim N\left(\frac{1}{K}\sum i' \in NE_i v_{i'}, \frac{1}{K\tau_v}\right)$, $i \neq i'$, where NE_i is the K-nearest neighbors of county i (K = 5) found by *knearneigh*() from R package *spdep,* and $\tau_v \sim Gamma(1, 0.001)$ is the conditional precision.

The global temporal effect $f_T(t) = \sum_{k=1}^{2} \delta_k C_k(t)$ is modeled using restricted cubic splines (RCS) [26] with 3 evenly spaced knots, where $C_k(t)$'s are the basis functions (including t itself) and $\delta_k \sim Normal(0, 0.001)$ are the coefficients. Similarly, the space–time interaction $f_{ST}(i, t) = \sum_{k=1}^{2} \delta_{i,k} C_k(t)$ represents the county-specific temporal trend modeled by RCS with 3 evenly spaced knots, where $\delta_{i,k} \sim Normal(0, \tau_{\delta,k}^{-1})$ and $\tau_{\delta,k} \sim Gamma(1, 0.001)$ for $k = 1, 2$. These flexible global and local temporal trends partially account for the transmission nature of the disease, that is, the propagation and decay of the epidemics.

All covariates were standardized according to their mean and SD. The posterior distributions of the model parameters in the Bayesian spatiotemporal models introduced in this section were estimated using Integrated Nested Laplace Approximation (INLA) [27]. In the INLA framework, a Gaussian Markov random field is assumed for latent effects, and Laplace approximation is used to calculate posterior marginals as an alternative to the sampling-based Markov chain Monte Carlo (MCMC) method in Bayesian inference. INLA was implemented using R package R-INLA. Due to strong correlations between the movement components, for both the single-lag and DLM models, we assessed 1 movement component at a time.

For work-related movement (the index available for the most counties), we compared Model (1) with a negative binomial model with only fixed effects based on criteria including deviance information criterion (DIC), Watanabe-Akaike information criterion (WAIC), and marginal log-likelihood. We also estimated COVID-19 incidence using Model (1) with work-related mobility and compared the estimated incidences with the observed ones. In addition, we validated our model with leave-1-week-out cross-validation based on the weekly DLM model for work-related movement. Specifically, we trained the model with the training data including all but 1 week and predicted COVID-19 cases for the left-out week. This process was repeated for the 14 weeks of the study period. Cross-validated R^2 was calculated by regressing the aggregated observed weekly COVID-19 cases over 1944 counties against the aggregated predicted weekly COVID-19 cases over 1944 counties. Furthermore, we used a prediction model including an additional autoregressive term in Model (1) to predict 1-week-ahead COVID-19 incidence per 100K population for the week of July 16, 2020. The details of the prediction model are in Section 2 of the Supplementary Data.

RESULTS

The final numbers of counties used for the analyses of each mobility index were 1944 for workplace, 1520 for retail and

recreation, 1398 for grocery and pharmacy, 1240 for residential community, 900 for transit, and 657 for parks. The counties included in each movement analysis were mapped with colors indicating the cumulative incidences of confirmed COVID-19 cases during the 28 weeks from March 11 to July 15, 2020 (Supplementary Figure 1). County-level characteristics were similar across counties included in the analyses for the 6 movement indices, except the counties used in the analysis of parks-related movement had higher population densities and higher proportions of Latinos compared with those used for other mobility indices (Supplementary Table 2).

Figure 1 presents the smoothed incidence and movement trend (except Parks) for all states included in the analyses based on the Loess smoothing method with span = 0.05 (ie, a smoothing window of 7 days). For the residential movement component, positive values imply longer duration of staying

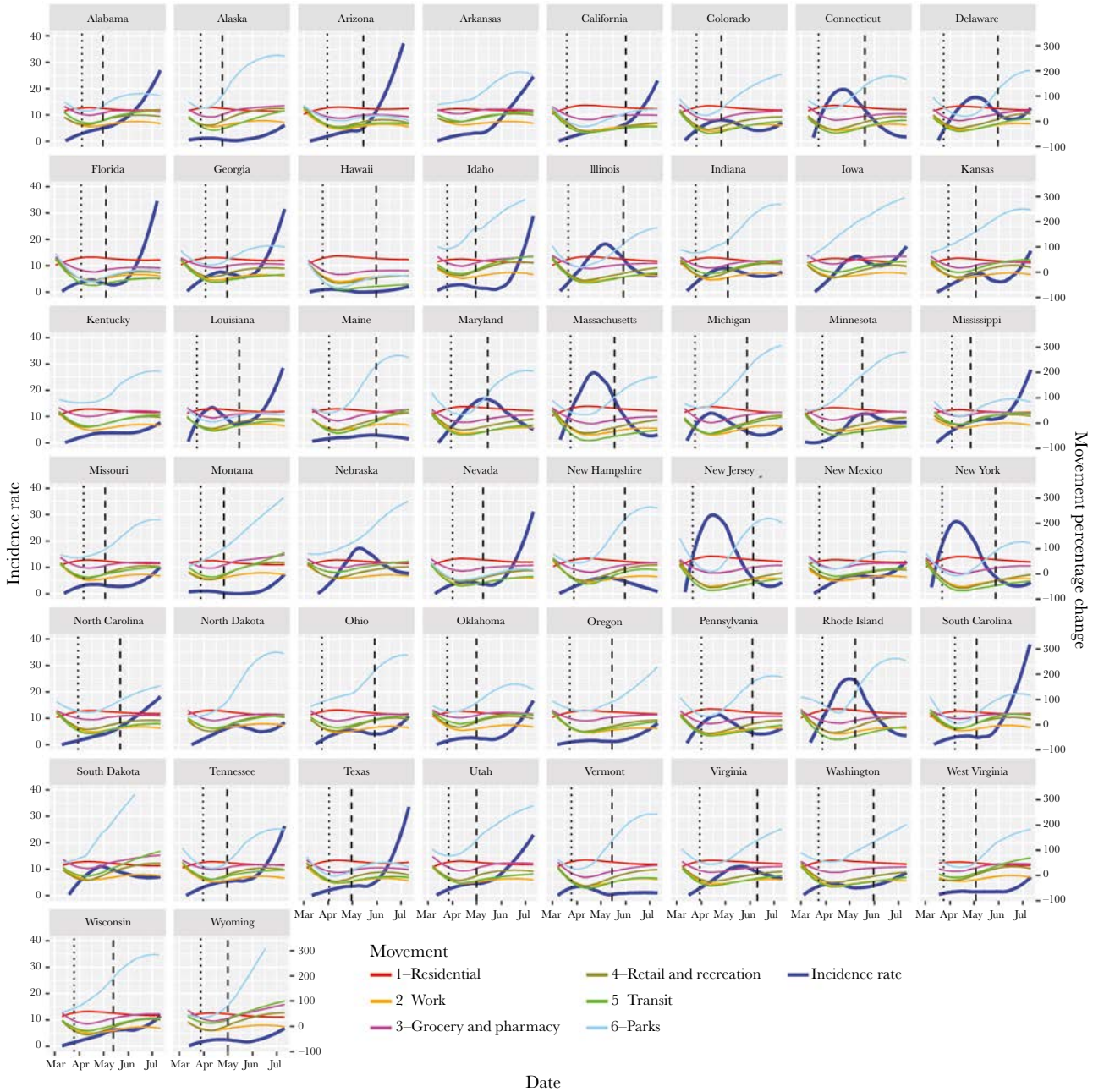

Figure 1. Smoothed incidence and movement trend (except Parks) for all states based on the Loess smoothing method with span = 0.05 (ie, a smoothing window of 7 days); the black dotted and dashed lines correspond to shelter-in-place and state-reopening effective dates, respectively. For the residential movement component, positive values imply longer duration of staying at home than prepandemic time, and larger values imply longer duration of staying at home. For the other movement components, positive values imply more visitors to places than prepandemic time, and larger values imply more visitors to places. Parks is included in Supplementary Figure 2 due to its large scale.

at home than in prepandemic time (ie, the median duration in the 5-week period from January 3 to February 6, 2020 [18]). Parks is included in Supplementary Figure 2 due to its large scale. The abrupt decline of all nonresidential mobility indices started earlier than the initiation of "shelter in place" in all states where such interventions were implemented (Figure 1; Supplementary Figure 2). In nearly all states, the decline of nonresidential movement overlapped with the ascent of the spring wave of reported COVID-19 cases. The exponential increase of the spring wave was clear in many states such as Connecticut, Delaware, Louisiana, Massachusetts, Michigan, New Jersey, New York, and Rhode Island. The actual exponential increase in the number of infections might have occurred 2 or more weeks earlier, given the average incubation period of 5–7 days for SARS-CoV-2 and the delay in laboratory confirmation and reporting during the spring wave. In most states, the epidemic curve either plateaued or sharply declined shortly after the trough of the nonresidential mobility curves, implying an effect of movement restrictions on mitigating the local epidemics. The nonresidential mobility curves started to rise again before the official reopening announcement in all states. Though not fully recovered to the prepandemic level, the mobility indices gradually increased from mid-April to the end of June, with a substantial rise during the summer wave in many states, for example, Alabama, Arizona, Florida, Georgia, Idaho, Louisiana, Mississippi, Nevada, South Carolina, Tennessee, and Texas. In general, workplace-related mobility recovered more slowly, whereas parks-related mobility recovered more rapidly than other nonresidential mobility types. The patterns of associations between incidence and mobility trends observed at the national level were qualitatively similar to those in most Southern states (Supplementary Figure 3). All mobility indices except for parks-related mobility were highly correlated with each other, with absolute Spearman correlation coefficients ranging 0.65 to 0.91 (Supplementary Figure 4). Not surprisingly, residential movement was negatively correlated with all nonresidential movements, and the highest negative correlation with residential stay was observed for work-related movement, followed by those related to retail and recreation. The data used to calculate the correlations were extracted from COVID-19 Community Mobility Reports from Google [18].

Figure 2 presents estimated nationwide IRRs based on models with only a single lag by movement category. The estimated IRRs associated with a 10% increase in movement were in the expected direction only for lags of 6 weeks or longer. Consequently, we report the associations at the 7-week lag as primary results. A 10% increase in a given movement index 7 weeks prior was associated with an estimated IRR of 0.87 (95% credible interval [CI], 0.80–0.94) for residential, 1.05 (95% CI, 1.01–1.09) for workplaces, 1.08 (95% CI, 1.04–1.12) for retail and recreation stores, 1.05 (95% CI, 1.01–1.09) for transit, and 1.05 (95% CI, 1.02–1.09) for grocery and pharmacies, corresponding to increases (reductions if negative) in disease incidence rate of –13%, 5%, 8%, 5%, and 5%, respectively. No statistically significant association was found for parks-related movement. There is a substantial amount of spatial heterogeneity in the delayed effects of movement indices on disease spread in terms of the magnitude of the effects and the length of the delays (Figure 3; Supplementary Figures 5–9). Although North and South Carolina showed significant reductions in incidence rate as early as 2 weeks after the increase in the duration of residential stay, similar reductions did not occur until 7–8 weeks after in most states (Figure 3). Interestingly, statistically significant reductions were established at week 3 and remained so through week 8 in Texas, Oklahoma, Florida, and Vermont. Similar patterns were seen for movements related to work, retail/recreation, and grocery/pharmacy (Supplementary Figures 5–7). A small number of states showed quick increases in the reported incidences 1–2 weeks after movement increased, but it took 7–8 weeks to see increased incidences in more than half of the states. Washington, Oregon, Montana, Wyoming, Wisconsin, Texas, Arkansas, Georgia, and North Carolina showed persistent positive associations between disease incidence and work-related movement for lags of 3–4 weeks and longer. Transit- and especially parks-related movements showed much weaker associations with disease incidence.

Figure 4 presents results from the weekly distributed lag time models. One movement component is considered in each model, adjusting for time course of the pandemic, spatial autocorrelation, population density in 5 quantiles, proportion of population in 3 age groups (20–45, 45–65, and >65), proportion of population with at least a Bachelor's degree, percentage of the population below the poverty line, percentage of African Americans, percentage of Hispanics, average winter/summer maximum temperature in degrees Fahrenheit, and percent obese. As shown in Figure 4, the estimated IRR curves over the lag weeks based on the DLM analyses show similar patterns as those seen in the single-lag analyses. Negative IRR estimates associated with residential duration were observed at a lag of 5 weeks and onwards. The tipping point of IRR estimates (from negative to positive values) occurred at a lag of 5 weeks for work-related and parks-related movements and at 6 weeks for movements related to retail and recreation, grocery and pharmacy, and transit. The magnitudes of estimated IRRs at a lag of 7 weeks were smaller than the ones based on single-lag models, as the effects were distributed among the multiple lags in the DLM.

For work-related movement, Model (1) has smaller values of DIC and WAIC and a larger value of marginal likelihood compared with the negative binomial model with fixed effects (Supplementary Table 3), showing that our mixed-effects model has a better fit than the model with only fixed effects. In addition, for the DLM model for work-related movement, the estimated and observed cumulative COVID-19 incidence rates

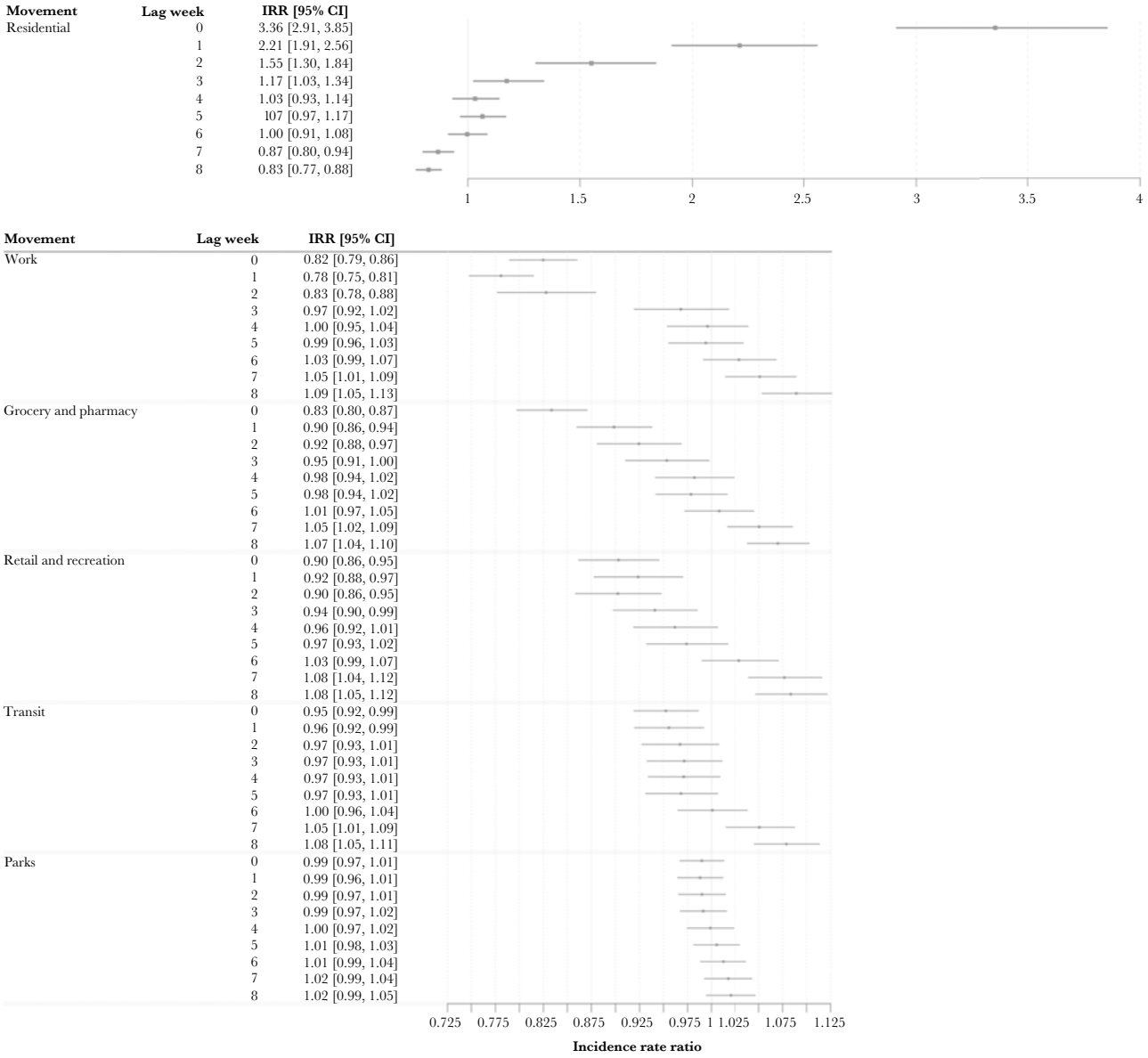

Figure 2. Analysis results from the weekly single-lag models. The estimated IRR for the fixed effect is per 10% increase in movement component u weeks ago ($u = 0, 1, 2, \ldots, 8$). One movement component with 1 single lag time is considered in each model, adjusting for time course of the pandemic, spatial autocorrelation, and population density expressed in 5 quantiles, proportion of population in 3 age groups (20–45, 45–65, and >65), proportion of population with at least a high school degree, poverty percentage, proportion of African Americans, proportion of Hispanics, average winter/summer maximum temperature in degrees Fahrenheit, and percent obese. Abbreviation: IRR, incidence rate ratio.

per 100K population across US counties during the week of July 9, 2020, had strong agreement (Supplementary Figure 10), indicating satisfactory goodness of fit. The estimated cumulative COVID-19 incidences per 100K population for all weeks of the study period in the 46 states combined are also highly consistent with the observed values (Supplementary Figure 11). In addition, the cross-validated R^2 was 0.86 from the leave-1-week-out cross-validation, indicating good model performance. Based on our prediction model for work-related movement, the predicted 1-week-ahead COVID-19 incidence for the week of July 16, 2020, agreed well with the observed values (Section 2 of the Supplementary Data).

DISCUSSION

Nearly 2 years since the first reported case in December 2019 [20], the COVID-19 pandemic is still threatening the world, with the new Delta variant dominating the spread [28]. While effective vaccines and therapies have become increasingly available, NPIs such as school closures [10], travel bans [11], business suspensions, social distancing [12], and increased testing capacity remain imperative for containing the spread of the disease. Most of these interventions directly constrain human movement. Using a Bayesian hierarchical spatiotemporal model, we assessed the effects of changes in multiple

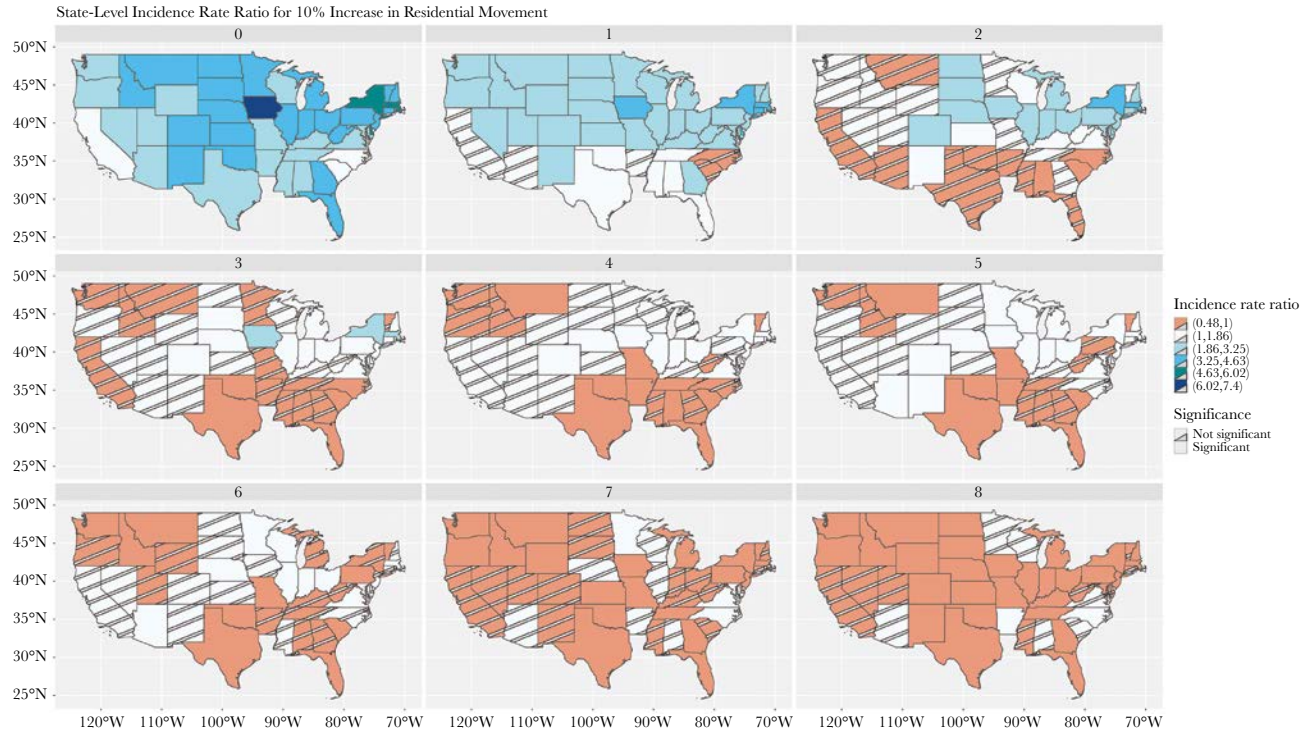

Figure 3. State-specific incidence rate ratio for 10% increase in residential mobility estimated by single-lag models, with warm colors showing a protective effect against COVID-19 spread and cold colors showing a harmful effect toward COVID-19 spread. Each map represents 1 specific lag week (from 0 weeks to 8 weeks). Shaded states indicate no statistical significance; that is, the 95% credible interval includes 1. Abbreviation: COVID-19, coronavirus disease 2019.

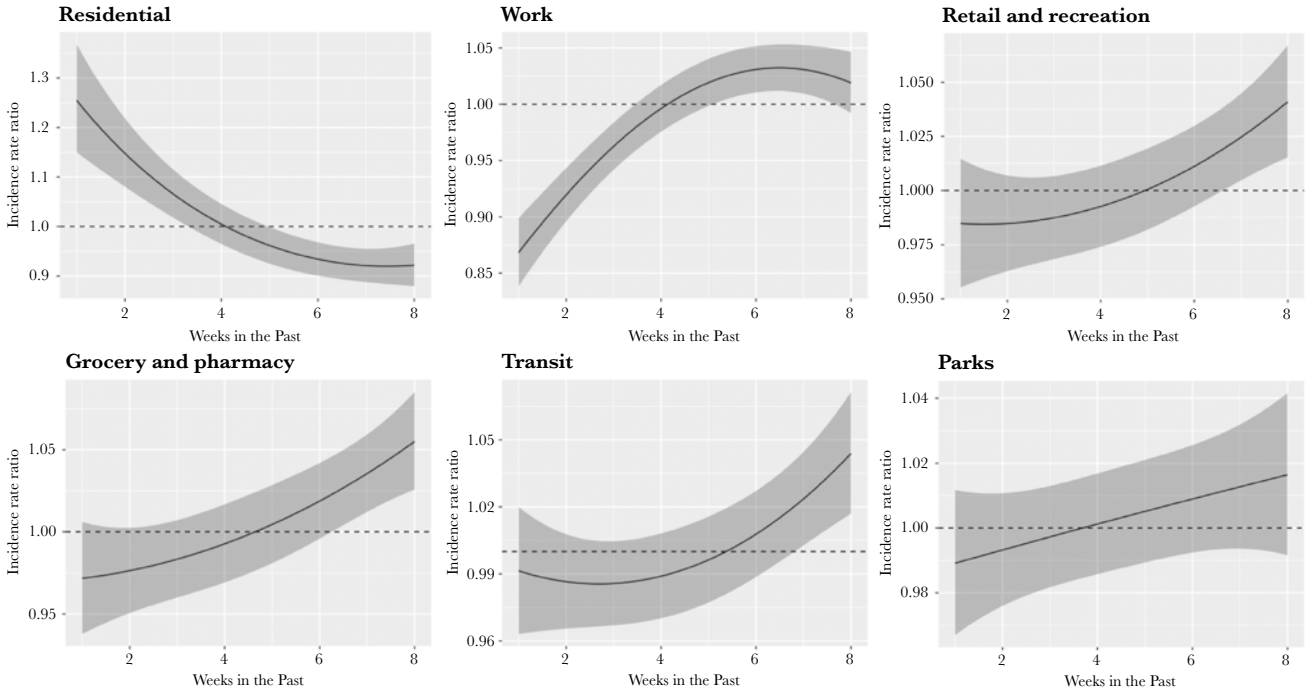

Figure 4. Results from the weekly distributed lag time models; the y-axis is the incidence rate ratio for the fixed effect by a 10% increase in movement component; the shaded areas represent 95% pointwise credible intervals. One movement component is considered in each model, adjusting for time course of the pandemic, spatial autocorrelation, population density in 5 quantiles, proportion of population in 3 age groups (20–45, 45–65, and >65), proportion of population with at least a high school degree, poverty percentage, proportion of African Americans, proportion of Hispanics, average winter/summer maximum temperature in degrees Fahrenheit, and percent obese.

types of human movement on reported incidence of COVID-19 and the potential delays of these effects in the United States. We found that time spent in residence was by far the most influential factor, with a 10% increase associated with a >10% reduction in incidence. The impacts of nonresidential movement types appeared comparable to each other, except for parks-related movement, which showed minimal impact on disease incidence. It took several weeks for the changes in movement to show significant impact on incidence in many states.

Human mobility is a key mediator between NPIs and the spread of COVID-19 [29], yet most previous studies [1–6, 10–14] on the effectiveness of NPIs in the United States did not examine the relationship between changes in human mobility and disease incidence, and thus their results are difficult to compare or even conflicting. For example, Courtemanche et al. concluded that the implementation of shelter-in-place led to a moderate reduction of 8.6% in the COVID-19 growth rate after 21 days, holding school closure, ban on large group gathering, and nonessential business closure constant [13]. Brauner et al., however, found only a small effect of the stay-at-home order conditioning on closing schools and universities, banning gatherings, and closing nonessential businesses [14]. Our study took advantage of the availability of human movement data captured by smart phones and quantified the association between changes in human mobility indices and the incidence of reported COVID-19 cases, providing valuable and unambiguous input for future modeling studies and policy-makers. For example, our IRR estimates of 1.05–1.08 for every 10% increase in nonresidential movements are translated into overall ratios of 1.22–1.36 (1.05^4–1.08^4) for a 40% increase. In other words, a 40% reduction in work-related travel, which was reached in many states (Figure 1), leads to an estimated 22%–36% reduction in the incidence rate. Likewise, the maximum increase of 20% in residential duration in many states corresponds to an estimated 26% reduction in the incidence rate.

Based on both single-lag and distributed-lag models, changes in movement start affecting the reported COVID-19 incidence with an estimated lag of 5–7 weeks. This range of lags is reasonable because of the incubation period of 5–14 days for COVID-19 [30], a possible delay of a few days from symptom onset to testing, and another 1–2 weeks to get the test results back according to a July 11th report by *CNET Health and Wellness* [30]. If taking into account this delay in testing and case reporting, the lag for movement to start affecting COVID-19 incidence might be around 3–6 weeks. The DLMs tended to give slightly shorter delays and smaller IRRs than the single-lag models, but overall the 2 approaches yielded similar estimates (Figure 2 vs Figure 4). This similarity confirms the robustness of our findings. The difference in the single-lag models and the DLMs is due in part to the autocorrelation among movements in consecutive weeks. In the DLM model, the use of smoothing splines alleviates but does not entirely remove the collinearity issue. Such collinearity may partially explain the counterintuitive inverse association between disease incidence and movement changes at short lags. Another explanation for the inverse association is that a high COVID-19 incidence rate could reduce nonresidential movements and thereby increase residential stay due to increased perception of risk. The single-lag models avoid the collinearity issue at the price of attributing all effects to a single lag, which likely overestimates the effect of the given lag. The 2 approaches should be viewed as complementary to each other and together provide a better picture about the true effects.

We found that both the delay and magnitude of the effects of movement changes on COVID-19 incidence varied greatly among states. We assessed the Spearman correlation between state-specific delay time and state-level variables for residential movement. The delay time in each state was the week when the estimated IRR for that state from the single-lag model started to be <1. We found that the delay time was positively correlated with state-level median household income (r = 0.41) and negatively correlated with winter temperature (r = –0.33) and the percentage of below–high school education (r = –0.40). These correlations are consistent with the pattern in Figure 3 that Northeastern states (New England) tended to have longer delays than Southern and Northwestern states. This pattern might be explained by the possibility that wealthier states with higher education levels were more likely to have better social distancing, mask wearing, and other non-mobility-related preventions, which could dilute the immediate effect of mobility changes and thus lead to a longer time for the epidemic to respond to changes in mobility. Whether non-mobility-related preventions truly modulate the effect of human mobility is worth future study. The heterogeneity in the delay and magnitude of the effects of the movement changes might also result from differences in social culture, environment, and policies related to interventions, testing, and reporting across states, and the underlying drivers warrant further future investigation.

One limitation of this study is the underdetection of COVID-19 cases, as not everyone with symptoms in the United States was tested, especially early in the first wave, and some test results could be false-negative due to improper sample collection or insufficient viral shedding [31]. Underdetection of cases in the early phase could bias our results toward the null. In addition, Google collects mobility data from phone users who have opted in to location history services [18], leading to a potentially biased sample population in the study and limiting the generalizability of our findings. For example, 58% of counties in the data set for workplace and 72%–86% of counties in the data sets for the other mobility components are among the top 40% most densely populated counties in the United States (Supplementary Table 2). However, the issue of representativeness likely exists in all types of phone-based or app-based mobility data. Moreover, we did not explicitly model disease transmission or the effects

of movement changes on transmission parameters, such as the effective reproductive number. Nevertheless, because the county-specific epidemic curves unexplained by human movement changes were captured by the county-specific spatiotemporal effect (the term $f_S(i) + f_T(t) + f_{ST}(i,t)$in the model) and the estimated effects are multiplicative factors on incidence rate, the estimated effects should be comparable to those derived from a transmission model. Finally, for short-term prediction, we incorporated an autoregressive term in the GAMM model, which reflects the transmission nature of COVID-19 and yielded reasonable predictive performance. Time series models, for example, the Bayesian structural time series model, could be an alternative for both estimation and prediction as they can handle complex autocorrelation structures [32], but the underlying assumption of Gaussian-distributed outcomes may not suit low-incidence settings at the county level.

While potent vaccines are more and more widely available, the inadequate supply leaves room for new SARS-CoV-2 variants of concern to surface. For the foreseeable future, NPIs, especially those limiting human movement, may continue to be the primary option in countries experiencing short supply of vaccines or the emergence of vaccine-resistant variants. When initiating movement restrictions, policy-makers should anticipate the possibility of delay of the effect by several weeks and should communicate this anticipation to the public to increase compliance. In countries with high vaccine coverages, a longer delay is possible if non-mobility-related preventions indeed modulate the effect of human mobility. The finding that movement associated with parks may not affect COVID transmission is encouraging, as outdoor activities can help us maintain a physically and socially active lifestyle without increasing the risk of infection. In the long run, NPIs will be the first option for newly emerging, highly contagious pathogens. Future research should prioritize methods of monitoring human movement in an efficient, representative, and ethical way to form a clearer picture about the effectiveness and timeliness of human movement in mediating the relationship between intervention policies and disease transmission. Policy-makers should anticipate the delay in the effectiveness of NPIs when planning intervention strategies.

Supplementary Data

Supplementary materials are available at *Open Forum Infectious Diseases* online. Consisting of data provided by the authors to benefit the reader, the posted materials are not copyedited and are the sole responsibility of the authors, so questions or comments should be addressed to the corresponding author.

Acknowledgments

Availability of data and code. Data used in this study have been made publicly available by various sources; please refer to Supplementary Table 1 for full links to access the data set. R code can be found at https://github.com/chloehe1129/COVID_spatiotemporal.

Author contributions. S.H. and J.L. were in charge of the literature search, data collection, visualization, data analysis, and manuscript writing. M.W. motivated the study design and research question. S.H. and J.L. wrote the first draft. M.W. and Y.Y. supervised data analyses and revised the manuscript to its final version, with input from P.J. and B.L. All authors contributed to data interpretation. J.X. contributed to data visualization.

Patient consent. This study does not include factors necessitating patient consent. The study only involves publicly accessible deidentified and aggregated data and is therefore exempt from IRB approval.***Financial support.*** Y.Y. was supported by grant R56 AI148284 from the US National Institutes of Health. P.J. was supported by grant R01 HL150119 from the US National Institutes of Health. M.W. was supported by grant R01 ES026246 from the US National Institute of Health.

Role of the funding source. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Potential conflicts of interest. All authors: no potential conflicts of interest. All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

References

1. Tian H, Liu Y, Li Y, et al. An investigation of transmission control measures during the first 50 days of the COVID-19 epidemic in China. Science **2020**; 368:638–42.
2. Lai S. Ruktanonchai NW, Zhou L et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature **2020**; 585:410–3.
3. Chinazzi M, Davis JT, Ajelli M, et al. The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. Science **2020**; 368:395–400.
4. Rowan, NJ, Moral RA. Disposable face masks and reusable face coverings as non-pharmaceutical interventions (NPIS) to prevent transmission of SARS-CoV-2 variants that cause coronavirus disease (COVID-19): role of new sustainable NPI design innovations and predictive mathematical modelling. Sci Total Environ **2021**; 772:145530.
5. Duhon J, Bragazzi N, Kong JD. The impact of non-pharmaceutical interventions, demographic, social, and climatic factors on the initial growth rate of COVID-19: a cross-country study. Sci Total Environ **2021**; 760:144325.
6. Chowdhury R, Heng K, Shawon MSR. et al. Dynamic interventions to control COVID-19 pandemic: a multivariate prediction modelling study comparing 16 worldwide countries. Eur J Epidemiol **2020**; 35:389–99.
7. Badr HS, Du H, Marshall M, et al. Association between mobility patterns and COVID-19 transmission in the USA: a mathematical modelling study. Lancet Infect Dis **2020**; 11:P1247–54.
8. Chang S, Pierson E, Koh PW, et al. Mobility network models of COVID-19 explain inequities and inform reopening. Nature **2021**; 589:82–7.
9. Xiong C, Hu S, Yang M, et al. Mobile device data reveal the dynamics in a positive relationship between human mobility and COVID-19 infections. Proc Natl Acad Sci U S A **2020**; 117:27087–9.
10. Auger KA, Shah SS, Richardson T, et al. Association between statewide school closure and COVID-19 incidence and mortality in the US. JAMA **2020**; 324:859–70.
11. Adekunle AI, Meehan M, Rojas D, et al. Delaying the COVID-19 epidemic in Australia: evaluating the effectiveness of international travel bans. Aus N Zeal J Public Health. **In press**.
12. Aslam F. COVID-19 and importance of social distancing. Preprints 2020040078 [Preprint]. 7 April **2020**. Available at: https://www.preprints.org/manuscript/202004.0078/v1. Accessed 7 December 2021.
13. Courtemanche C, Garuccio J, Le A, et al. Strong social distancing measures in the United States reduced the COVID-19 growth rate. Health Aff. **In press**.
14. Brauner JM, Mindermann S, Sharma M, et al. Inferring the effectiveness of government interventions against COVID-19. Science **2021**; 371:eabd9338.
15. Brassey J, Heneghan C, Mahtani KR, Aronson JK. Do weather conditions influence the transmission of the coronavirus (SARS-CoV-2). The Centre for Evidence-Based Medicine: Evidence Service to Support the COVID-19 Response. 23 March **2020**. Available at: https://www.cebm.net/covid-19/do-weather-conditions-influence-the-transmission-of-the-coronavirus-sars-cov-2/. Accessed 7 December 2021.
16. DiMaggio C, Klein M, Berry C, Frangos S. Black/African American communities are at highest risk of COVID-19: spatial modeling of New York City ZIP code–level testing results. Ann Epidemiol **2020**; 51:7–13.
17. Statista.com. Most popular mapping apps in the United States as of April 2018, by monthly users. **2018**. Available at: https://www.statista.com/statistics/865413/

most-popular-us-mapping-apps-ranked-by-audience/. Accessed 7 December 2021.
18. Google LLC. Google COVID-19 community mobility reports. Available at: https://www.google.com/covid19/mobility/. Accessed 28 August 2021.
19. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis **2020**; 20:533–4.
20. Centers for Disease Control and Prevention. About CDC COVID-19 data: confirmed & probable counts. **2020**. Available at: https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/about-us-cases-deaths.html. Accessed 13 July 2020.
21. Stolberg SG. Trump Administration Strips C.D.C. of Control of Coronavirus Data. *New York Times*. 14 July **2020**. Available at: https://www.nytimes.com/2020/07/14/us/politics/trump-cdc-coronavirus.html. Accessed 15 July 2020.
22. Wu X, Nethery RC, Sabath MB, et al. Air pollution and COVID-19 mortality in the United States: strengths and limitations of an ecological regression analysis. Sci Adv **2020**; 6:eabd4049.
23. Zanobetti A. Generalized additive distributed lag models: quantifying mortality displacement. Biostatistics **2000**; 1:279–92.
24. Blangiardo M, Cameletti M, Baio G, Rue H. Spatial and spatio-temporal models with R-INLA. Spat Spatio-temporal epidemiol **2013**; 4:33–49.
25. Besag J, York J, Mollié A. A Bayesian image restoration with two applications in spatial statistics. Ann Inst Statist Math **1991**; 43:1–20.
26. Durrleman S, Simon R. Flexible regression models with cubic splines. Stat Med **1989**; 8:551–61.
27. Rue H, Martino S, Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Stat Methodol **2009**; 71:319–92.
28. Kupferschmidt K, Wadman M. Delta variant triggers new phase in the pandemic. Science **2021**; 372:1375–6.
29. Askitas N, Tatsiramos K, Verheyden B. Estimating worldwide effects of non-pharmaceutical interventions on COVID-19 incidence and population mobility patterns using a multiple-event study. Sci Rep **2021**; 11:1972.
30. Lauer SA, Grantz KH, Bi Q, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Ann Intern Med **2020**; 172:577–82.
31. Centers for Disease Control and Prevention. CDC 2019-novel coronavirus (2019-nCoV) real-time RT-PCR diagnostic panel. Available at: https://www.fda.gov/media/134922/download. Accessed 7 December 2021.
32. Feroze N. Forecasting the patterns of COVID-19 and causal impacts of lockdown in top five affected countries using Bayesian structural time series models. Chaos Solitons Fractals **2020**; 140:110196.