UC Irvine
UC Irvine Previously Published Works

Title
An online advertising intervention to increase adherence to stay-at-home-orders during the COVID-19 pandemic: An efficacy trial monitoring individual-level mobility data

Permalink
https://escholarship.org/uc/item/09q057gf

Authors
Garett, Renee R
Yang, Jiannan
Zhang, Qingpeng
et al.

Publication Date
2022-04-01

DOI
10.1016/j.jag.2022.102752

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

ELSEVIER

Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.

International Journal of Applied Earth Observations and Geoinformation 108 (2022) 102752

Contents lists available at ScienceDirect

International Journal of Applied Earth Observations and Geoinformation

journal homepage: www.elsevier.com/locate/jag

An online advertising intervention to increase adherence to stay-at-home-orders during the COVID-19 pandemic: An efficacy trial monitoring individual-level mobility data

Renee R. Garett [a,1], Jiannan Yang [b,1], Qingpeng Zhang [b], Sean D. Young [c,d,*]

[a] ElevateU, Irvine, CA, USA
[b] School of Data Science, City University of Hong Kong, Hong Kong, China
[c] Department of Emergency Medicine, University of California, Irvine, CA, USA
[d] University of California Institute for Prediction Technology, Department of Informatics, University of California, Irvine

ARTICLE INFO

Keywords:
Digital health
Intervention
COVID-19
Mobility
Artificial intelligence

ABSTRACT

The COVID-19 pandemic has led public health departments to issue several orders and recommendations to reduce COVID-19-related morbidity and mortality. However, for various reasons, including lack of ability to sufficiently monitor and influence behavior change, adherence to these health orders and recommendations has been suboptimal. Starting April 29, 2020, during the initial stay-at-home orders issued by various state governors, we conducted an intervention that sent online website and mobile application advertisements to people's mobile phones to encourage them to adhere to stay-at-home orders. Adherence to stay-at-home orders was monitored using individual-level cell phone mobility data, from April 29, 2020 through May 10, 2020. Mobile devices across 5 regions in the United States were randomly-assigned to either receive advertisements from our research team advising them to stay at home to stay safe (intervention group) or standard advertisements from other advertisers (control group). Compared to control group devices that received only standard corporate advertisements (i.e., did not receive public health advertisements to stay at home), the (intervention group) devices that received public health advertisements to stay at home demonstrated objectively-measured increased adherence to stay at home (i.e., smaller radius of gyration, average travel distance, and larger stay-at-home ratios). Results suggest that 1) it is feasible to use mobility data to assess efficacy of an online advertising intervention, and 2) online advertisements are a potentially effective method for increasing adherence to government/public health stay-at-home orders.

1. Introduction

The COVID-19 pandemic has been the most disastrous global health problem within the past one hundred years. According to the World Health Organization, as of March 16th, 2022, there have been more than 458 million cases of COVID-19 worldwide, including more than 6 million deaths. Within the United States (US), there have been almost 78 million cases and 960,703 deaths (WHO, 2022). A number of government policy mandates and recommendations have occurred in response to the pandemic and needs for controlling it. For example, in March of 2020, when the pandemic was just beginning in the United States, many US state governors issued stay-at-home-orders, including governors of California, Illinois, Texas, New York, and Florida, requiring (all but essential working) individuals within their states to remain at home to reduce viral transmission (Moreland et al., 2020). New interventions are needed to prevent and control COVID-19 as well as future pandemics.

A variety of new and existing technologies and tools are being created and/or applied to address the COVID-19 pandemic that might lead to interventions (Budd et al., 2020; Kumar et al., 2020; Mbunge et al., 2021; Young et al., 2022). For example, studies using mobility data have surged during the pandemic due to opportunities for public–private partnerships that were not possible before the pandemic (Buckee et al., 2020). These studies use global position system (GPS) traces and other mobility metric data (taken from smartphone data

* Corresponding author at: University of California Institute for Prediction Technology, Department of Informatics, 6091 Bren Hall, University of California Irvine, USA.

E-mail address: syoung5@hs.uci.edu (S.D. Young).

[1] Joint first authors.

https://doi.org/10.1016/j.jag.2022.102752
Received 20 October 2021; Received in revised form 17 March 2022; Accepted 19 March 2022
Available online 24 March 2022
156-8432/© 2022 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

when users move and these moves are recorded) to align with COVID-19-related outcomes. COVID-19 mobility studies have found that people's mobility patterns can monitor and predict COVID-19 transmission, (Gatalo et al., 2021; Wang and Yamamoto, 2020) adherence to stay-at-home-orders, (Jeffrey et al., 2020; Moreland et al., 2020) political views on COVID-19-related behaviors, (Clinton et al., 2021) and other important epidemiological and policy-related outcomes. Mobility data are therefore a new and growing data source that might inform COVID-19 and other infectious disease prevention and control efforts (Garett and Young, 2021).

However, one of the limitations in recent studies on mobility data is that mobility data are typically provided by companies in aggregate (i.e., data on the total number of visits to a location) rather than device-level (i.e., identifying the specific locations visited for each unique device). Individual-level interventions are typically more effective compared to broader (less personalized) interventions because they can be tailored to individual characteristics (Bull et al., 1999). Although insights from studies using aggregated mobility data provide important insights, individual-level data studies are needed both to better understand the context of an individual's mobility as well as to provide more granular data to inform future potential policy efforts. Individual-level mobility data might therefore improve the likelihood that health departments and researchers can act on human mobility data and apply this information in interventions.

Similarly, there have been a large number of studies and interventions using digital advertisements and social media to prevent and control COVID-19 and other infectious diseases (Huang et al., 2016; Young et al., 2022). By using different platforms that allow delivery of online advertisements, such as Facebook, Google, and other independent advertising platforms, researchers and health departments have the potential to rapidly scale outreach efforts to increase people's adherence to health recommendations, including for prevention, mask-wearing, and vaccination (Young and Schneider, 2020). Singapore has partnered with the app, WhatsApp, to send COVID-19-related information and government updates to citizens (Ting et al., 2020). The US Centers for Disease Control and Prevention (CDC) has partnered with advertising companies to attempt to reach out people at risk with online advertisements (Adams, 2021; AP-NEWS, 2020). Researchers have also used social media, such as Facebook, to recruit participants affected by the COVID-19 pandemic and connect them with online interventions (Singh et al., 2020). Importantly, these methods using online advertisements have the ability to incorporate methods from social psychology to improve the effectiveness of the advertising/behavior change, such as combining social normative theories into the messaging (Cialdini et al., 2006; Young and Goldstein, 2021).

A study in this area would be novel for a number of reasons. First, limited research has studied whether sending people online advertisements to change their behaviora actually leads to behavior change. Second, no known research studies whether targeting people with online advertisements based on their mobility can increase adherence to public health recommendations or orders. Finally, current studies using mobility data are typically incorporating aggregate (rather than individual-level) mobility data, which as described above, has limitations. This manuscript seeks to conduct a novel study to address those areas: specifically, we seek to send people online advertisements to stay at home based on their individual-level mobility patterns, and continue to monitor their mobility patterns to learn at a granular whether and how devices have responded to advertisements advising them to stay-at-home.

We focused the intervention on the regions of Chicago, Miami, New York, Dallas, and Greater Los Angeles/Orange County to gain a diverse demographic data sample across the US and because citizens within these states had been ordered by their governors to stay at home for all but essential activities.

2. Material and methods

2.1. Study design

2.1.1. Sampling

We developed and sent city/region-level polygons for each of the 5 regions above to our advertising partner, Cphere, who then provided us a with a list of the most frequently listed points of interest (POIs) for each region that was visited by mobile devices in their dataset for the month of April 2020, up through April 24th. Their dataset includes approximately 30 million active monthly device identification numbers (ID numbers used to identify smartphones) in the United States, along with latitude and longitude coordinates of the locations they had visited. Data on the locations/visits of these devices are only available if users 1) downloaded a mobile application that requests their device ID-level mobility data, 2) agreed to the terms and conditions of that mobility in tracking their mobility, 3) and the app provided their data to Cphere and/or a corporate affiliate of Cphere that shares these data.

We filtered that list of POIs to provide Cphere with 25 POIs of locations that were not health, education/school, grocery stores, or other potentially essential locations to attempt to limit the data to devices that had recently visited non-essential locations for each of the 5 regions. We chose to use 25 locations because we expected this would provide us with a sufficient number of devices who had visited these locations and allow us the time/effort to manually review all POI's within the location for being non-essential locations. Specifically, the process was as follows: 1) Sort POI visits from greatest to least number of visits; 2) For each POI, conduct manual google maps and other related searches to identify the characteristics of the POI to manually classify it as being an essential or non-essential location (we have described that this is not fool-proof; it's very possible that some of these locations either were incorrectly labeled, or changed their status due to changing policies during the course of the study, however, we believe this is not a major flaw (more on this below); 3) Identify the top 25 most visited non-essential POI's for each of the 5 locations, leading to a total of 125 POIs (5*125 POIs). To arrive at this process of obtaining the list of 125 POI's, approximately 1000 POI's had to be manually reviewed on google searches. We chose the top 25 most visited locations in each region, *excluding* healthcare, education, and other locations we thought would have visitors for only essential activities. For example, the list included bars and churches if they were in the top 25 most visited locations within the city, but we excluded hospitals even if they were in the top 25 most visited locations as people typically visit hospitals for essential needs. The types of POI's differed based on city/region, as different cities/regions had different locations that were most visited.

It was not possible to rule out the possibility that devices were visiting these locations for an essential activity (e.g., an essential worker at a restaurant) for various reasons, including changes in policy/definition of essential during the course of the study. However, the identification of these "non-essential" POI's was intended to identify a group of devices that would be likely to be mobile during the course of the study to allow sufficient variation in mobility patterns for potential differences due to the intervention condition. Using this list of "non-essential POIs," Cphere identified 52,000 randomly-selected devices that had visited these POIs within the past 2 weeks distributed across the 5 regions (New York, n = 15,073; LA/OC, n = 992; Dallas, n = 15,428; Chicago, n = 12,506; Miami, n = 8494).

2.1.2. Randomizing

Randomization was a key component of these methods: The study was designed so that half of the devices (control group devices/No-Ad group) would only see typical advertisements that they or any other device would normally see that are paid by corporate brands (e.g., Southwest Airlines, McDonalds, or local shoe stores). The other half of devices (intervention group devices) would see these same ads but instead of some of those ads, they would see our advertisements advising them to stay at home. In other words, an approximately equal number of devices from Chicago (or one of the other regions, as we stratified by region) would either see traditional corporate ads (control/No-Ad group) or corporate ads *and* our public health ads (intervention group). The study was designed this way to ensure that devices would either see our public health stay-at-home ads, or not. They were randomly assigned to one of these 2 groups only after all devices were collected for each region to improve randomization methodology. The final sample included the following number of intervention devices, which was approximately half of the total sample due to random assignment (New York, n = 7,537; LA/OC, n = 496; Dallas, n = 7,714; Chicago, n = 6,253; and Miami, n = 4,247.

2.1.3. Intervention

For the intervention group devices, Cphere was given advertisements focused on the importance of adhering to stay-at-home recommendations. Three types of advertisements were created: 1) pure informational (e,g., "Prevent the spread of the virus, stay at home"), 2) social norms (e.g., "everyone is staying home"), and 3) punitive (e.g., if you don't stay at home you could get fined). If users clicked on the ads, all ads directed them to a website at predictiontechnology.ucla.edu that provided more details on the importance of staying at home to prevent the spread of the virus. Both intervention and control/No-Ad devices only received an advertisement (whether it was a stay-at-home ad for the intervention group, or a standard ad for the control group) if and when they visited/viewed a mobile app on their device that belonged to the network of potential corporate affiliate apps that Cphere was able to post advertisements and receive location data (e.g., the Weather app). The intervention official began on April 29th, 2021 and last for the next 7 days. Cphere was instructed to show the stay-at-home advertisements to intervention devices up to 625,000 times (impressions). Any additional ads over this amount that were shown to intervention group devices would be regular corporate advertisements, similar to the control group. The other half of devices (control/No-Ad devices) did not receive the public health advertisements from our research team related to stay-at-home-orders, but instead received only the regular corporate advertisements received by visitors of mobile apps (e.g., advertisements to purchase shoes, etc).

2.2. Mobile phone location data

The location data of the devices were collected by Cphere throughout the study period, from baseline (before April 29th, 2021) to post-7-day-intervention, including a few additional days of follow-up (through May 10th, 2021). Similar to Cphere's ability to place advertisements, Cphere was only able to collect mobility data from devices if and when those devices visited an app that shares mobility data with Cphere and the user agreed to provide these data. The data are anonymized, thus no individual can be specified, and personal information such as gender, age, and occupation is unknown. Each GPS record contains the unique ID, timestamp, longitude, and latitude of the device. The data acquisition frequency of GPS locations changes according to the device's movement speed to minimize the burden on the device's battery. If it is determined that the user is staying somewhere for a long time, the data is acquired at a relatively low frequency, and if it is determined that the user is moving, the data is acquired more frequently.

2.3. Data preprocessing

Due to the weak GPS signal or low battery of devices, location data may have contained invalid data. Similar to other studies using methods attempting to filter out invalid self-report survey data, GPS, and/or social media data, (Winter et al., 2019; Young, 2012; Young et al., 2020) we developed methods to preprocess the data to improve data quality. In this paper, the data cleaning for our mobile phone location data is based on three assumptions. First, we assumed the first location of each device was the correct one since we could not know which record is correct. Second, a certain amount of distance shifting (1 km) was allowed. Third, abnormal moving speed (larger than 160 km/h) between two consecutive data records was determined to be unacceptable and removed from the dataset. The distribution of the processed data across the whole study period is shown in Fig. 1.

For each device, suppose the first two consecutive data records are location A and B. Based on our assumptions, location A is assumed as a correct record. Location B was determined to be invalid if satisfying two conditions: 1) the distance between A and B is larger than 1 km, 2) and the moving speed from A to B is larger than 160 km/h. We repeated this process for all the records of each device. For example, suppose one device has 4 records in one day, say point A, B, C, D, and the distance between these points are $d_{AB} = 4\text{km}$, $d_{BC} = 0.5\text{km}$, $d_{CD} = 0.1\text{km}$, with the time duration between these points $t_{AB} = 72\text{s}$, $t_{BC} = 60\text{s}$, $t_{CD} = 1\text{s}$. Under our assumptions, we assumed that the first data point is correct, which means the record of point A is acceptable, and after computing the speed from A to B, says $s_{AB} = 200\text{km/h}$, which exceeds the upper limit of the normal moving speed (160 km/h) and we will delete point B. For the next record C, we compute the speed between A and C, which is $s_{AC} = \frac{d_{AC}}{t_{AC}} = 100\text{km/h}$, thus we assume point C is acceptable. For point D, the speed between C and D $s_{CD} = 360\text{km/h}$, but since the distance between C and D is within a tolerable shifting distance (1 km) (usually due to GPS shifting), we still assume point D is correct.

2.4. Determining home location from trajectory data

The home location for each device was determined by applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) (Ester et al., 1996) algorithm to the nighttime stay points (observed between 8p.m. and 6 a.m.) of each device. Each location is represented by its geographic coordinates (longitude and latitude), and the inputs of the DBSCAN algorithm are a series of locations observed during nighttime. DBSCAN algorithm can find several cluster centers by high density and expands clusters from them. Specifically, to determine one cluster, DBSCAN algorithm will first randomly choose one location as the cluster center and retrieve this location's neighborhood locations within ε distance far, and if this cluster contains sufficiently many points (*minPts*), a cluster is founded. Since multiple cluster centers may be generated by the DBSCAN algorithm, we determine the home location as the cluster center has the largest number of data points. In this paper, ε is set to 1 km, and *minPts* is 10. The implementation of DBSCAN is based on scikit-learn packages on Python and the distance and neighborhood was determined by haversine formula and ball tree algorithm, respectively, where haversine formula is used to compute the Euclidean distance based on geographic coordinates and ball tree algorithm is used for spatial division of data points and their allocation into certain clusters. For more details of the DBSCAN algorithm, please refer to the Supplementary Information.

2.5. Individual mobility indices

Given a sequence of GPS locations $P_i = p_i^1, p_i^2, \cdots, p_i^N\}$ in a single day of one device *i*, we apply three indices to characterize individual mobility patterns:

Radius of Gyration (R_g) (Gonzalez et al., 2008) is used to measure the depth of a trajectory, that is how far the entity went.

$$R_g = \frac{1}{N}\sum_{n=1}^{N} dist\left(p_i^n, \overline{p}_i\right), \quad (1)$$

Where $\overline{p}_i$ is the center of the sequence and $dist()$ denotes the distance measure.

Average travel distance (*ATD*) (Yabe et al., 2020) is the average of great circle distance (geographic distance) between all subsequent pairs of GPS observations on a given day.

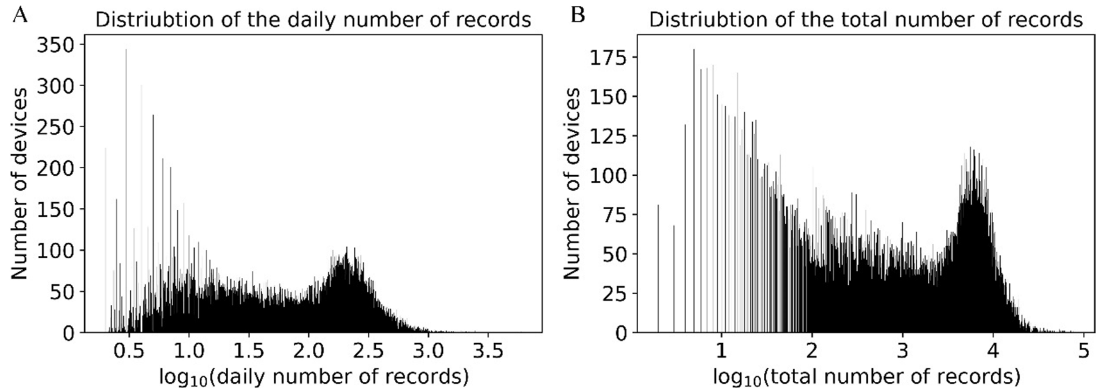

Fig. 1. Data distribution in view of the daily number of records (A) and the total number of records (B) throughout the whole study period. Note the x-axis of the two plots are both the logarithmic value of the corresponding number of records with a base of 10.

$$ATD = \frac{1}{N}\sum_{n=1}^{N} dist\left(p_i^n, p_i^{n+1}\right), \tag{2}$$

Stay-at-home ratio within a spatial threshold k (SAH_k) (Yabe et al., 2020) is the ratio of individual device who had stayed the entire day within a distance k from the estimated home location.

$$SAH_k = \frac{1}{N}\sum_{n=1}^{N} I\{dist\left(p_i^n, p_i^H\right) \leq k\}, \tag{3}$$

Where $I\{\}$ is the indicator function and p_i^H is the home location of device i, which is determined by DBSCAN algorithm.

2.6. Difference-in-Differences (DiD) analysis

We adopted a DiD analysis (Dimick and Ryan, 2014) to measure the effects of the devices receiving only, versus receiving and clicking on the advertisements. It was important to distinguish between these behaviors as clicking on an ad is a more active behavior compared to just receiving the ad, and it is possible that users of devices receiving ads did not actually see the received ad (e.g., they might have quickly moved past the ad). Specifically, the observations are divided into two groups: control/No-ad group and treatment/intervention group, where treatment indicates the groups receiving and clicking advertisements. A regression model is formulated to estimate the average treatment effects.

$$y_{ijt} = \beta_0 + \beta_1 treat_j + \beta_2 post_{jt} + \delta\left(treat_j \times post_{jt}\right) + \varepsilon, \tag{4}$$

Where y_{ijt} denotes the value of one specific individual mobility index i (such as radius of gyration or average travel distance) for one device j at time t, $\beta_0, \beta_1, \beta_2, \delta$ are all parameters in this regression model, and ε denotes the random error term. In Equation (4), $treat_j$ and $post_{jt}$ are both dummy variables, where $treat_j$ denotes whether device j has specific treatment (received or received and clicked on the ad), and $post_{jt}$ indicates whether device j has received treatment by time t, and $post_{jt} = 0$ if time t before treatment and $post_{jt} = 1$ if time t after treatment. Thus, $treat_j \times post_{jt}$ is an interaction term, where $treat_j \times post_{jt} = 1$ if one device j has already received treatment at time t, and $treat_j \times post_{jt} = 0$ otherwise. The parameter δ indicates the average treatment effect of one specific treatment (Ryan et al., 2015). Note that the statistical analysis of δ is implemented by two-tailed test.

2.7. Statistical analysis

We sought to identify the change in these mobility indices (i.e., the radius of gyration, stay-at-home ratio (within 500 m, 1000 m and 3000 m), and average travel distance) from baseline (April 27th, 2021) to post-intervention (May 10th, 2021) between control/No-ad and intervention (including received and clicked) groups. For the devices in different groups (such as the control/No-ad devices that did not receive any public health advertisements and the devices that received and also clicked on the advertisements), we utilized the Kolmogorov Smirnov test (KS-test) to check whether two groups follow the same distribution and compare their mean values to see whether the difference is statistically significant. Because there are many more (control/No-ad) devices that did not receive public health advertisements than those that received and clicked, we utilized the Bootstrap sampling method to estimate both the mean and median values of mobility indexes for the devices that did not receive advertisement to make a more balanced comparison. P-values of $p < .05$ were generated to determine significant differences in these mobility indices.

3. Results

542 of the devices in the intervention group received at least one COVID-19 stay-at-home advertisement. 468 intervention devices only received these advertisements without any further operations (e.g., didn't click on them), which are denoted as Received. 74 intervention devices received and clicked on the advertisements; these devices are represented as Clicked. For the control/No-ad devices that did not receive any public health advertisements and only received standard corporate advertisements, we denote them as No-AD.

Fig. 2 and Figure S1 show the mean and median values between April 27th, 2020 (before advertisements were sent) and May 10th, 2020 (after the intervention was completed) of the radius of gyration (R_g), average travel distance (ATD), and stay-at-home ratio within 500 m (SAH_{500}) and stay-at-home ratio within 3000 m (SAH_{3000}), respectively. Note that the mean values of the No-AD group shown in Fig. 2 (blue solid curves) are generated by the Bootstrap sampling method and the confidence interval can be found in Fig. 3. As shown in both Fig. 2 and Figure S1, there are obvious differences between the devices that did not receive our public health advertisements (No-AD, control group devices, blue solid curves), the devices that only received our public health advertisements (Received, green solid curves), and the devices that both received and also clicked on our public health advertisements (Clicked, red solid curves) in view of all the individual mobility indices. Specifically, compared to devices that didn't receive any advertisements, the devices that received advertisements (Received and Clicked) had a statistically significant smaller radius of gyration, average travel distance, and larger stay-at-home ratios. To be more intuitive, we chose one day (May 7th, 2020) as an example to show the differences between these groups which is shown in Table 1. The individual mobile activity of the devices in Received group is between the ones of No-Ad group and the Clicked group. There were no statistical differences between the devices that only received advertisements (Received) andthe devices that received and also clicked on the advertisements (Clicked) SAH_{500}, SAH_{1000}, SAH_{3000}) Table 2.

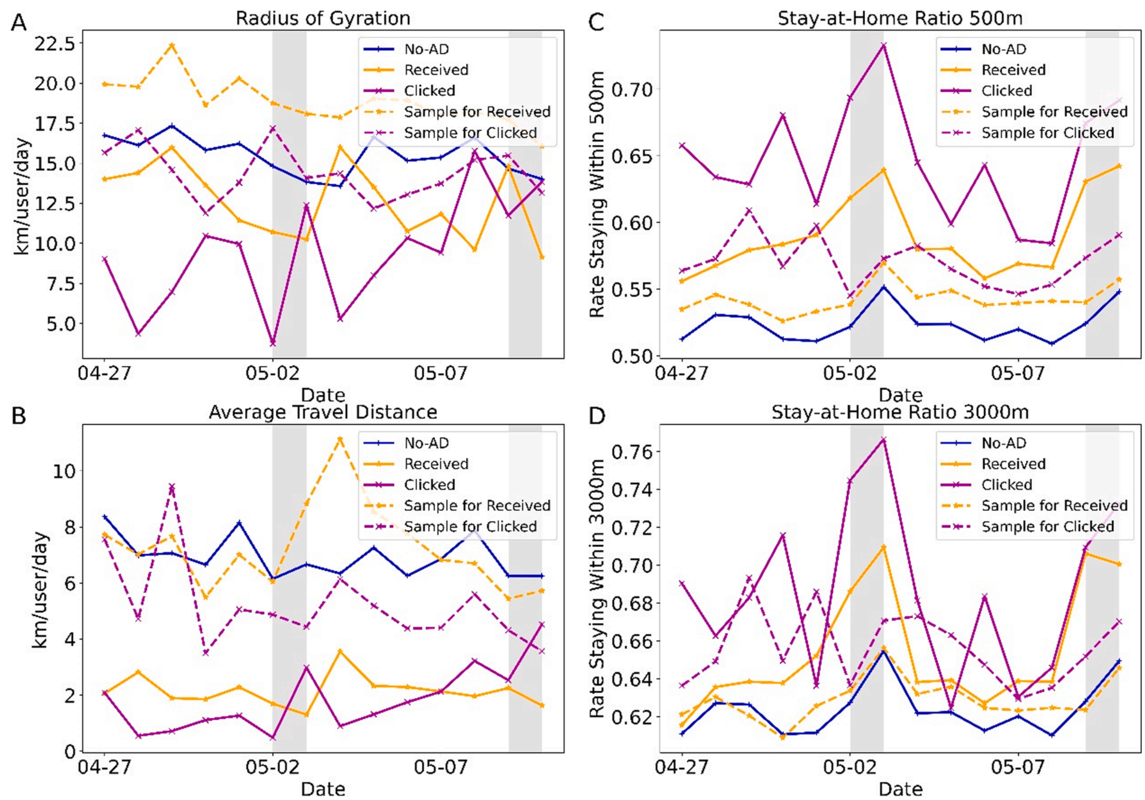

Fig. 2. Changes in individual mobility patterns. Temporal transition of mean values in view of radius of gyration (A), average travel distance (C), stay-at-home ratio with a 500 m threshold (B) and 3000 m threshold (D). The colors denote the devices in different states: the devices in the (control/No-ad) group that only received standard corporate ads but not public health (intervention group) ads (No-AD) (blue), the intervention group devices that received public health ads but didn't click on them in Received (orange), and the intervention group devices that received and clicked on the public health ads in Clicked (purple). The solid and dashed curves in this figure represent the original and sampled data, respectively. The grey bars denote the weekends. Note that the mean values of the No-AD group are generated by the Bootstrap sampling method and the confidence interval can be found in Fig. 3. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

To verify whether the difference in lower individual mobility index for the devices who received advertisements (Received and Clicked) was caused by their unique data distributions (for example, more/fewer data records, or outliers), we randomly sampled an equal number of devices for the Received and Clicked devices from No-AD devices, respectively. Specifically, for each device in Received and Clicked, the sampled device in No-AD group must satisfy the requirement that it has the most similar number of daily records to reduce the possibility that differences were due to devices with a larger or smaller number of daily records. The sampling results are shown as dashed curves in Fig. 2 and Figure S1. As we can see, after only including devices with a similar number of records, the devices in Clicked still showed a significantly smaller radius of gyration (p-value < 0.0001), total travel distance (p-value < 0.001), and larger stay-at-home ratios (p-value < 0.0001 in view of SAH_{500}) compared with the corresponding Clicked sampling devices. Due to a small number of overall clicks, we did not conduct analyses analyzing differences in click-through-rates by specific message type.

Further, DiD analysis revealed that all the average treatment effects are in line with the aforementioned findings. However, due to the small number of devices that eventually received and clicked the advertisements, the effects were typically not significant, except for the reduced *ATD* as a result of receiving the advertisements.

4. Discussion

Results suggest the feasibility and efficacy of using online advertising to increase adherence to government-mandated stay-at-home orders, along with use of mobility data to track behavior change resulting from online advertising. As a result of the lifting of stay-at-home orders during the intervention period, mobility metrics increased among both the intervention and control groups. However, mobility remained lower among the intervention group devices (especially those that not just received a public health ad but also clicked on it) compared to those (control group devices) that did not receive them, suggesting efficacy of the intervention in reducing mobility and increasing adherence to stay-at-home orders. The results generated in DiD analysis are in line with these findings. The additional analyses conducted among Received and Clicked devices with a similar number of records confirmed these results. These additional analyses suggest that the reduction in mobility during COVID-19 among these devices that received advertisements was not caused by outliers in these data attributes, but by willingness of these device owners to adhere to stay-at-home orders.

There are several immediate actionable insights from this work that can be taken by researchers and health departments. First, researchers and health departments should make use of online advertising methods for public health behavior change, especially at the early stages of a public health emergency, by exploring the potential of placing advertisements on people's mobile phones that encourage people to mitigate risk behaviors and improve health promotion. In addition, similar to studies that have shown that social media, internet search, and other digital data can be used to remotely monitor health behaviors and outcomes (Young et al., 2019), cell phone mobility data should be explored as a method for near real-time monitoring of public health behavior change. Finally, given the privacy issues in monitoring using mobile devices, researchers should explore the ongoing ethical and implementation science questions associated with these approaches.

We initially found that the control group/No-ad devices had a larger stay-at-home ratio but a similar distance traveled. To explore and

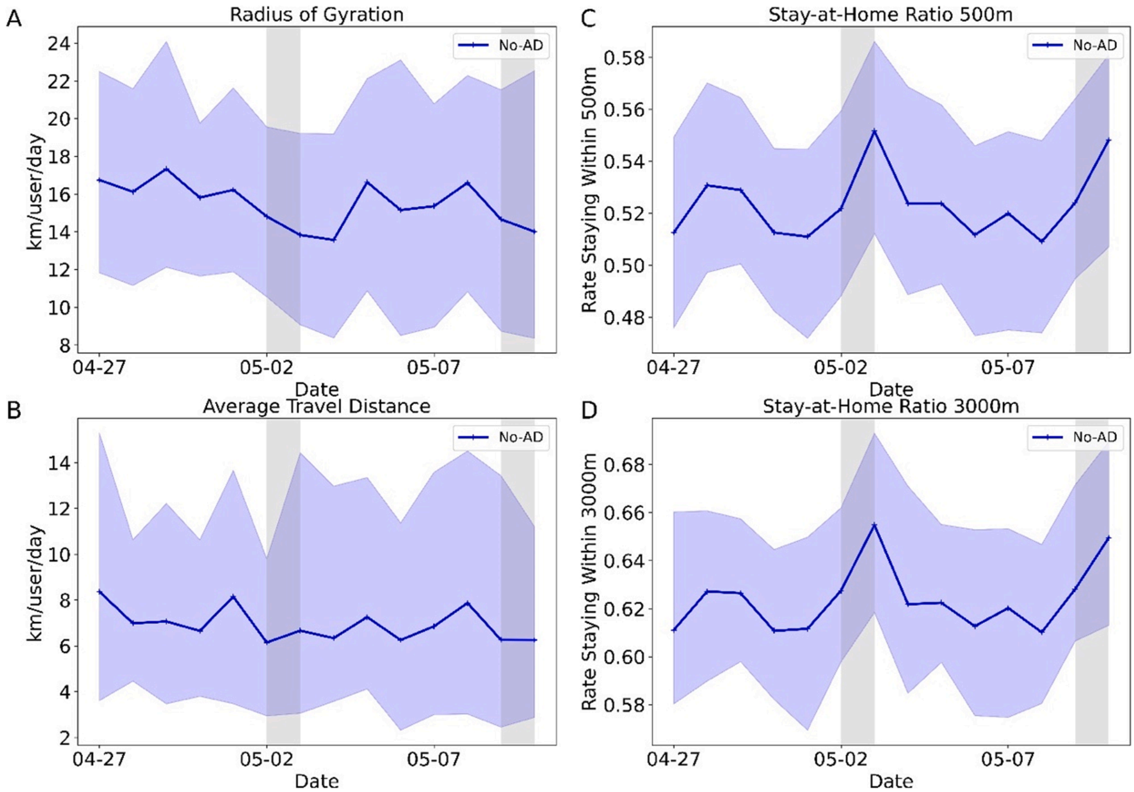

Fig. 3. Bootstrap estimated mean values of four individual mobility indexes for No-AD/control groups. (A). Radius of gyration. (B). Average travel distance. (C). Stay-at-home ratio within 500 m. (D). Stay-at-home ratio within 3000 m. The shadow areas in the subplots denote the 95% confidence intervals.

Table 1
Statistics of different groups in view of different individual mobility indices on 05/07, 2020 (Thursday). Received, Clicked and No-AD/control group ads in this table denotes the devices in the intervention group that received public health ads but did not click the advertisements, the devices in the intervention group that received public health ads and clicked on the advertisements, and the devices in the control/no public health advertisement group that did not receive public health stay-at-home advertisements and only received standard ads, respectively. The p-values are derived by KS-test by comparing the observations of different groups. * denotes a significant difference between two observations.

Statistics	R_g(km)	ATD(km)	SAH_{500}	SAH_{1000}	SAH_{3000}
Mean of No-AD	15.80	6.76	0.52	0.55	0.61
Mean of Received	13.22	3.24	0.56	0.58	0.63
Mean of Clicked	9.11	2.06	0.58	0.59	0.64
p-value between Received and No-AD	<0.0001*	<0.0001*	<0.0001*	<0.0001*	<0.0001*
p-value between Clicked and No-AD	0.001*	<0.001*	0.007*	0.002*	<0.001*
p-value between Received and Clicked	0.513	0.49	0.82	0.56	0.66

* Significant at p <.05.

address this issue, we set the home location for each device using the DBSCAN algorithm, and then assumed that the area within 1 km radius was also home (to account for the GPS shifting). Therefore, the possible explanation for the larger stay-at-home ratio but similar distance traveled could be that the users had outdoor activities within a 1 km radius.

Table 2
The estimated average treatment effects in view of different individual mobility indexes. The values in the brackets denote the p-values of parameter δ generated by two-tailed test. In view of the radius of gyration (R_g) and average travel distance (ATD), the negative δ s suggest a trend in the direction that receiving the public health advertisements would reduce these two indexes, with a statistically significant difference such that the devices that received the public health (intervention group) ads had a lower travel distance than those that only received the standard ads and not the public health ads (No-Ad/control group). In contrast, the positive δ s given three stay-at-home ratios suggest that devices that received and/or clicked on advertisements (i.e., any devices in the intervention group who were sent public health ads) were trending (not significant) in the direction of being more likely to stay at home.

Treatment	Groups	δ_{R_g}	δ_{ATD}	$\delta_{SAH_{500}}$	$\delta_{SAH_{1000}}$	$\delta_{SAH_{3000}}$
Receiving	Received vs No-AD	−1.53 (0.52)	−0.85 (0.02*)	0.02 (0.38)	0.01 (0.55)	0.02 (0.30)
Clicking	Clicked vs No-AD	−3.82 (0.62)	−0.08 (0.94)	0.12 (0.19)	0.13 (0.14)	0.13 (0.13)
Clicking	Clicked vs Received	−4.03 (0.61)	−1.33 (0.46)	0.11 (0.30)	0.12 (0.21)	0.11 (0.26)

* Significant at p <.05.

For example, people may have wanted to exercise close to their home by going for short walks to get out of the house but staying in very close proximity.

Although early and requiring additional research beyond this pilot prior to implementation, these methods and results are already of potentially high impact to the field of public health and related intervention research. For example, related to the methods, no known research has explored the integration of online advertisements and mobility data, making this an important first study. In addition, although there has been a lot of work on the use of online advertisements in marketing, including for public health, these studies have typically focused on measuring more immediate responses to the advertisements, such as “clicks,” website visits, or purchases on a website. The fact that we were able to link online advertisements to reductions in mobility that may have occurred more than a week later provide a unique benefit to the field of marketing and public health on potential methods that can be used to measure the impact of online advertisements. When adding the finding that these advertisements reduced people’s mobility (an important behavioral and public health outcome), it further strengthens the potential impact of the study by demonstrating an easy method of intervention to improve adherence to COVID-19 and other future public health recommendations. In fact, as mobility and other smartphone data continue to grow and become accessible to researchers, studies may increasingly begin to use these digital footprints to measure changes in people’s behaviors as a result of interventions. Finally, although this study focused on adherence to stay-at-home orders, a policy which is no longer in place in most regions, the methods and findings could easily be applied to other government recommendations. For example, as tracking and reducing mobility remains an important issue to prevent the spread of COVID-19 and other/future infectious diseases, these online intervention and mobility monitoring methods might be suitable tools to be studied in the future and integrated into public health efforts.

It is worth noting that LA/OC had a smaller number of enrolled devices that had visited the 25 POIs compared to other locales, even LA/OC has a large population. We are unable to know why we found this result. One hypothesis is that we chose a large number of POI’s from LA/OC that were along the beach because we expected these would be considered non-essential locations. However, during the time of the study the state required beaches to shut down, making it possible that very few individuals visited these specific POIs.

This study is limited by a number of factors. First, we were unable to gain personal information about the device users which may influence their views on COVID-19 and willingness to adhere to government recommendations, such as their race, age, ethnicity, political affiliation, etc. However, this limitation is largely offset by having a control group of randomly assigned devices to not receive advertisements. Similarly, we do not have information on the devices who received advertisements and didn’t click on them. Having this additional information would help us to understand potential biases within the study and may also help us to better understand the role of the online advertisements in reducing mobility and to understand how to plan future interventions to get people to seek more public health information by clicking on advertisements. We also did not have data on the number of advertisements received by each device. As devices that received more advertisements would be more likely to actually view and click on those ads, it will be important for future research attempting to better understand this approach seek to address those questions. We were also limited by our reliance on a corporate partner who was bound to non-disclosure and sales agreements preventing them from sharing more detailed information about device owners (e.g., demographic information). As this is a first study exploring these types of methods, we encourage future research to consider these issues with data sharing agreements to explore more research questions about the sample (e.g., political ideology, age, etc) (Grossman et al., 2020). We were also limited by the short time frame of the study, designed to react quickly to the pandemic, but not allowing us to follow-up with devices longitudinally. We were also limited by the quality of data provided by the provider, including accuracy of the location data. Future research can better explore the accuracy of data provided by advertising technology companies. The study was also limited by the sample size of devices that received advertisements, including the small size of the Received group and Clicked group. Although the results generated by DiD analysis are in line with exploratory findings, these results are not statistically significant, likely due to the limited devices in Received and Clicked groups. This is further supported as the one statistically significant difference was found among the device groups with the largest number of devices (received vs no ad groups). In our study, to receive an advertisement, during the study period, devices had to open an app that was accessible for our ad technology partner to send an advertisement and had to have their setting checked to continue receiving ads and having their mobility data tracked. Although devices had these settings in place when they were recruited, it is possible that they changed their privacy settings during the weeks of the study or did not open an eligible app, meaning that they would not have been able to receive an advertisement/notification. Future research can further explore how to best work with advertising partners to send digital public health and safety outreach communications to devices in ways that will be seen and have the most reach, as well as implementation science research on how to appropriately conduct large-scale public health research like this in real-world settings. For example, more studies are needed to explore how to track face mask wearing and vaccination uptake in real world settings.

5. Conclusion

New methods are needed to help reduce COVID-19 and other infectious diseases transmission. This study suggests that it is feasible and potentially effective to use online advertisements to increase people’s adherence to public health behavior change orders (i.e., staying at home and/or reducing mobility during COVID-19) and to use mobility data to monitor adherence. As technologies, modeling methods, and technology data continues to increase, health departments will increasingly seek to integrate and optimize their existing methods (e.g., online health promotion outreach) by integrating it with novel data and modeling methods, such as mobility data.

CRediT authorship contribution statement

Renee R. Garett: Conceptualization, Methodology. **Jiannan Yang:** Data curation, Analysis. **Qingpeng Zhang:** Investigation, Methodology, Analysis, Supervision. **Sean D. Young:** Conceptualization, Methodology, Funding.

Declaration of Competing Interest

SDY is a Co-investigator on a small business grant application to incorporate mobility methods in public health and is an advisor to startups on this topic. All other authors have no conliects to report.

Acknowledgments

This work was funded by the National Institute on Allergy and Infectious Diseases (Young 7R01MH106415), National Institute on Drug Abuse (NIDA), and National Center on Complementary and Integrative Health (NCCIH).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jag.2022.102752.

References

Adams, K., 2021. “CDC, Ad Council launch campaign to combat COVID-19 vaccine hesitancy with support from major brands.” https://www.beckershospitalreview.com/digital-marketing/cdc-ad-council-launch-campaign-to-combat-covid-19-vaccine-hesitancy-with-support-from-major-brands.html (accessed May 27, 2021).

AP-NEWS, 2020. “Leading Digital and Tech Platforms Join Ad Council, White House, HHS and CDC to Address COVID-19 Crisis,” AP NEWS, Mar. 25, 2020. https://apnews.com/press-release/PR%20Newswire/60cc6ff4fae333d280a1d261b858f88a (accessed Dec. 16, 2020).

Buckee, C.O., Balsari, S., Chan, J., Crosas, M., Dominici, F., Gasser, U., Grad, Y.H., Grenfell, B., Halloran, M.E., Kraemer, M.U., 2020. Aggregated mobility data could help fight COVID-19. Science.

Budd, J., Miller, B.S., Manning, E.M., Lampos, V., Zhuang, M., Edelstein, M., Rees, G., Emery, V.C., Stevens, M.M., Keegan, N., 2020. Digital technologies in the public-health response to COVID-19. Nat. Med. 26, 1183–1192.

Bull, F.C., Kreuter, M.W., Scharff, D.P., 1999. Effects of tailored, personalized and general health messages on physical activity. Patient Educ. Couns. 36, 181–192.
Cialdini, R.B., Demaine, L.J., Sagarin, B.J., Barrett, D.W., Rhoads, K., Winter, P.L., 2006. Managing social norms for persuasive impact. Social influence 1, 3–15.
Clinton, J., Cohen, J., Lapinski, J., Trussler, M., 2021. Partisan pandemic: How partisanship and public health concerns affect individuals' social mobility during COVID-19. Sci. Adv. 7, eabd7204.
Dimick, J.B., Ryan, A.M., 2014. Methods for evaluating changes in health care policy: the difference-in-differences approach. JAMA 312, 2401–2402.
Ester, M., Kriegel, H.-P., Sander, J., Xu, X., 1996. A density-based algorithm for discovering clusters in large spatial databases with noise. Kdd 226–231.
Garett, R., Young, S.D., 2021. Geolocation, ethics, and HIV research. Health Technol. 11, 1305–1309. https://doi.org/10.1007/s12553-021-00611-0.
Gatalo, O., Tseng, K., Hamilton, A., Lin, G., Klein, E., 2021. Associations between phone mobility data and COVID-19 cases. Lancet. Infect. Dis 21, e111.
Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.-L., 2008. Understanding individual human mobility patterns. Nature 453, 779–782.
Grossman, G., Kim, S., Rexer, J.M., Thirumurthy, H., 2020. Political partisanship influences behavioral responses to governors' recommendations for COVID-19 prevention in the United States. Proc. Natl. Acad. Sci. 117, 24144–24153.
Huang, E., Marlin, R.W., Young, S.D., Medline, A., Klausner, J.D., 2016. Using Grindr, a smartphone social-networking application, to increase HIV self-testing among Black and Latino men who have sex with men in Los Angeles, 2014. AIDS Educ. Prev. 28, 341–350.
Jeffrey, B., Walters, C.E., Ainslie, K.E., Eales, O., Ciavarella, C., Bhatia, S., Hayes, S., Baguelin, M., Boonyasiri, A., Brazeau, N.F., 2020. Anonymised and aggregated crowd level mobility data from mobile phones suggests that initial compliance with COVID-19 social distancing interventions was high and geographically consistent across the UK. Wellcome Open Research 5.
Kumar, A., Gupta, P.K., Srivastava, A., 2020. A review of modern technologies for tackling COVID-19 pandemic. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14, 569–573.
Mbunge, E., Akinnuwesi, B., Fashoto, S.G., Metfula, A.S., Mashwama, P., 2021. A critical review of emerging technologies for tackling COVID-19 pandemic. Human behavior and emerging technologies 3, 25–39.
Moreland, A., Herlihy, C., Tynan, M.A., Sunshine, G., McCord, R.F., Hilton, C., Poovey, J., Werner, A.K., Jones, C.D., Fulmer, E.B., 2020. Timing of state and territorial COVID-19 stay-at-home orders and changes in population movement—United States, March 1–May 31, 2020. Morb. Mortal. Wkly Rep. 69, 1198.
Ryan, A.M., Burgess Jr, J.F., Dimick, J.B., 2015. Why we should not be indifferent to specification choices for difference-in-differences. Health Serv. Res. 50, 1211–1235.
Singh, P., Cumberland, W.G., Ugarte, D., Bruckner, T.-A., Young, S.D., 2020. Association between generalized anxiety disorder scores and online activity among US adults during the COVID-19 pandemic: cross-sectional analysis. Journal of medical Internet research 22, e21490.
Ting, D.S.W., Carin, L., Dzau, V., Wong, T.Y., 2020. Digital technology and COVID-19. Nat. Med. 26, 459–461.
Wang, H., Yamamoto, N., 2020. Using a partial differential equation with Google Mobility data to predict COVID-19 in Arizona. arXiv preprint arXiv:2006.16928.
Who, 2022. WHO Coronavirus (COVID-19) Dashboard. accessed Jun. 01, 2021. https://covid19.who.int.
Winter, N., Burleigh, T., Kennedy, R., Clifford, S., 2019. A simplified protocol to screen out VPS and international respondents using Qualtrics. Available at SSRN 3327274.
Yabe, T., Tsubouchi, K., Fujiwara, N., Wada, T., Sekimoto, Y., Ukkusuri, S.V., 2020. Non-compulsory measures sufficiently reduced human mobility in Tokyo during the COVID-19 epidemic. Sci. Rep. 10, 1–9.
Young, S.D., 2012. Recommended guidelines on using social networking technologies for HIV prevention research. AIDS Behav. 16, 1743–1745.
Young, S.D., Goldstein, N.J., 2021. Applying social norms interventions to increase adherence to COVID-19 prevention and control guidelines. Prev. Med. 145, 106424.
Young, S.D., Schneider, J., 2020. Clinical care, research, and telehealth services in the era of social distancing to mitigate COVID-19. AIDS Behav. 24, 2000–2002.
Young, S.D., Cumberland, W.G., Singh, P., Coates, T., Cumberland, W.G., 2022. A Peer-Led Online Community to Increase HIV Self-Testing among African American and Latinx MSM: A Randomized Controlled Trial. J. Acquired Immune Deficien. Syndrom. (1999). https://doi.org/10.1097/qai.0000000000002919. PMID: 35044989.
Young, S.D., Lee, S.-J., Perez, H., Gill, N., Gelberg, L., Heinzerling, K., 2020. Social media as an emerging tool for reducing prescription opioid misuse risk factors. Heliyon 6, e03471.
Young, S.D., Zhang, Q., Zeng, D.D., Zhan, Y., 2022. Cumberland W Social Media Images as an Emerging Tool to Monitor Adherence to COVID-19 Public Health Guidelines: Content Analysis. J Med Internet Res 24 (3), e24787. https://doi.org/10.2196/24787. PMID: 34995205.