Predicting spatial and temporal responses to non-pharmaceutical interventions on COVID-19 growth rates across 58 counties in New York State: A prospective event-based modeling study on county-level sociological predictors

Yunyu Xiao

Submitted to: JMIR Public Health and Surveillance
on: July 16, 2020

Table of Contents

Predicting spatial and temporal responses to non-pharmaceutical interventions on COVID-19 growth rates across 58 counties in New York State: A prospective event-based modeling study on county-level sociological predictors

Yunyu Xiao[1] PhD, MPHIL

[1]School of Social Work Indiana University–Purdue University Indianapolis Indianapolis US

Corresponding Author:
Yunyu Xiao PhD, MPHIL
School of Social Work
Indiana University–Purdue University Indianapolis
902 W. New York Street
Education/Social Work Building, ES 4119
Indianapolis
US

Abstract

Background: Non-pharmaceutical interventions (NPIs) have been implemented in the New York State since the COVID-19 outbreak on March 1, 2020 to control the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Projecting the growth rate of incidence as a response to key NPIs is crucial to guide future policy making. Few studies, however, considered spatial variations of incidence growth rate across different time points of NPIs.

Objective: This study quantifies county-level predictors of the time evolution of COVID-19 incidence growth rate following key NPIs in New York State.

Methods: County-level COVID-19 incidence data were retrieved from the Coronavirus Case Data from Social Explorer Website between March and June 2020. 5-day moving average growth rates of COVID-19 were calculated for 16 selected time points on the dates of eight NPIs and their respective 14-day-lag-behind time points. A total of 36 county-level predictors were extracted from multiple public datasets. Geospatial mapping was used to display the spatial heterogeneity of county-level COVID-19 outbreak. Generalized mixed effect least absolute shrinkage and selection operator (LASSO) regression was employed to identify significant county-level predictors related to the change of county-level COVID-19 growth rate over time.

Results: Since March 1, the growth rate of COVID-19 infection increased and peaked by the end of March, followed by a decrease. Over time, the region with the highest growth rates shifted from New York metropolitan area towards Western and Northern areas. Proportions of population aged 45 years and above (?=3.25 [0.17–6.32]), living alone at residential houses (?=3.31 [0.39–-6.22]), and proportion of crowd residential houses (?=6.15 [2.15–10.14]) were positively associated with the growth rate of COVID-19 infection. In contrast, living alone at rental houses (?=-2.47 [-4.83–-0.12]) and rate of mental health providers (?=-1.11 [-1.95–-0.28]) were negatively associated with COVID-19 growth rate across all 16 time points.

Conclusions: Tailored interventions and policies are required to effectively control the epidemic for different counties. Attention towards economic, racial/ethnic, and healthcare resource disparities are needed to narrow the unequal health impact on vulnerable populations.

(JMIR Preprints 16/07/2020:22578)
DOI: https://doi.org/10.2196/preprints.22578

Preprint Settings

1) Would you like to publish your submitted manuscript as preprint?
Please make my preprint PDF available to anyone at any time (recommended).
Please make my preprint PDF available only to logged-in users; I understand that my title and abstract will remain visible to all users.

Only make the preprint title and abstract visible.

✔ **No, I do not wish to publish my submitted manuscript as a preprint.**

2) If accepted for publication in a JMIR journal, would you like the PDF to be visible to the public?

✔ **Yes, please make my accepted manuscript PDF available to anyone at any time (Recommended).**

Yes, but please make my accepted manuscript PDF available only to logged-in users; I understand that the title and abstract will remain v

Yes, but only make the title and abstract visible (see Important note, above). I understand that if I later pay to participate in <a href="http

Original Manuscript

Predicting spatial and temporal responses to non-pharmaceutical interventions on COVID-19 growth rates across 58 counties in New York State: A prospective event-based modeling study on county-level sociological predictors

Yunyu Xiao PhD [ab*]

Affiliations:
[a] School of Social Work, Indiana University–Bloomington, Indiana, USA
[b] School of Social Work, Indiana University–Purdue University Indianapolis, Indiana, USA

***Corresponding Author**:
Yunyu Xiao, PhD, School of Social Work, Indiana University–Purdue University Indianapolis, IU School of Social Work, ES 4119, 902 W. New York St., Indianapolis, IN 46202, 2012537264; Email: yx18@iu.edu; Phone: +1 (201)-253-7264

Abstract

Background

Non-pharmaceutical interventions (NPIs) have been implemented in the New York State since the COVID-19 outbreak on March 1, 2020 to control the transmission of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). Socioeconomic heterogeneity across counties closely manifests differences in the post-NPIs growth rate of incidence, which is a crucial indicator to guide future infectious control policy making. Few studies, however, examined the geospatial and sociological variations in the epidemic growth across different time points of NPIs.

Objectives

To guide a more effective reopening plan while controlling the transmission, the current study aims at 1) identifying hotspots of the growth rate of COVID-19 incidence among the 57 counties and New York City in NYS over time, and 2) examining the association of COVID-19 growth rates after eight critical NPIs time points and most relevant county-level sociological predictors.

Methods

County-level COVID-19 incidence rates were retrieved from the Social Explorer Website between March 7, 2020 to June 22, 2020. 5-day moving average growth rates of COVID-19 incidence were calculated for 16 selected time points, including the dates of eight NPIs and their respective 14-day-lag-behind time points. A total of 36 county-level indicators were extracted from multiple public datasets. Geospatial mapping and heatmap were used to analyze spatial and temporal heterogeneity of county-level COVID-19 outbreak over selected NPIs-related dates. Generalized mixed effect least absolute shrinkage and selection operator (LASSO) regression, controlling for the 5-day moving average growth rates of COVID-19 testing rates, was employed to identify significant county-level predictors related to the changes of county-level COVID-19 growth rates over time.

Results

COVID-19 infection increased and peaked by the end of March (η=22.50%). Growth rates of COVID-19 decreased by 50.48% after implementing NPIs such as closures of schools, non-essential businesses, parks, and subways. There was a geospatial shift in the region with the highest growth rates from New York metropolitan area towards Western and Northern regions over time. Proportions of population aged 45 years and above (β=3.25 [0.17–6.32]), living alone at residential houses (β=3.31 [0.39–-6.22]), and proportion of crowd residential houses (β=6.15 [2.15–10.14]) were positively associated with the growth rate of COVID-19 infection. In contrast, living alone at rental houses (β=-2.47 [-

4.83–-0.12]) and rate of mental health providers (β=-1.11 [-1.95–-0.28]) were negatively associated with COVID-19 growth rate across all 16 time points.

Conclusions

There are geospatial differences in COVID-19 incidence after implementing different NPIs. Socioeconomic, racial/ethnic, and healthcare resource disparities at the structural and historical levels across counties need to be considered in infection control policymaking to narrow the unequal health impact on vulnerable populations effectively.

Keywords: Non-pharmaceutical interventions; COVID-19; incidence; growth rate; spatial; New York.

Introduction

Since the first confirmed cases of COVID-19 in New York State (NYS) on March 1, 2020, the statewide coronavirus outbreak has surged quickly. As of July 5, 2020, COVID-19 had affected more than 401 thousand individuals and caused nearly 32 000 deaths in NYS [1]. The State government has been actively responded to the transmission through a series of non-pharmaceutical interventions (NPIs) spanning March 2020 through late June 2020. Eight important NPIs were implemented in NYS, included declaring State of Emergency (March 7, 2020), school closure (March 16, 2020), NYS on PAUSE (March 20, 2020), park closure (April 1, 2020), acceptance of ventilator donation and USNS Comfort Hospital Ship (April 4, 2020), subway service suspension (April 25, 2020), and Phase I Reopening (June 8, 2020).

Statewide NPIs have a meaningful impact on the COVID-19 outbreak across NYS, reducing the number of cases by 94.10% from the peak date (April 14, 2020) as of June 8, 2020. Several modeling studies have revealed epidemiological variations in the impact of NPIs on COVID-19 incidence in China [2-4], UK [5], and Italy [6]. Yet, the impact of NPIs is not uniform across neighborhoods and communities, further contributing to health disparities in risk of COVID-19 infection rates in NYS [7-9]. No study, however, has empirically documented the effect of key statewide NPIs on the growth rate of COVID-19 incidence in different counties of NYS, the former epicenter of the COVID-19 pandemic in the United States (US).

Focusing on intra-state heterogeneity and temporal differences in response to NPIs has valuable public health importance. Cross-county variation in the epidemic growth reflects social processes, structural inequality, racism, intersectionality, and health disparities [10-13]. Since socioeconomically challenged communities/counties may experience a higher burden of disease and COVID-19 incidence [14], analyzing the spatial heterogeneity improves our understanding of the burdens of vulnerable communities and sociodemographic characteristics of the residents [15-17]. Such knowledge is important to guide tailored infection control policy actions, which may be more effective in reducing the COVID-19 growth, narrowing the gaps across geographic location and sociodemographic subgroups and promoting more socially just public health interventions [14, 17-21].

Addressing the county-level factors can provide more accurate information related to spatial and temporal variations. First, neighborhood effects on differences in health outcomes are complex and could be related to not only political-intuitional boundaries but micro-level differences segregated into county levels [13, 20]. COVID-19 growth rates may differ at the micro-level boundaries (e.g., county) but not macro-level boundaries (e.g., city, region), and *vice versa*. Thus, analyzing larger-scale boundaries may mask the heterogeneity within large areas. Second, it is possible that post-

NPI incidence rates are more reflective for residents with similar socioeconomic and demographic characteristics that drive the spatial segregation. In such a case, the study results shall be generalizable to other states or geographic districts in other countries based on the featured socioeconomic indicators. It is also plausible that geographic variations (e.g., availability and accessibility to health facilities, parks, recreation centers) are more relevant in COVID-19 incidence in response to different NPIs. Without much evidence from prior studies, the study does not begin with assumptions to preclude county-level variations that potentially reflect both geospatial and socioeconomic predictors. Third, most of the Census data and administrative data (e.g., insurance from the 2016 US Health, crime rates from the US Crime data, FBI) are collected at the county levels. This offers the opportunity to link multiple datasets to examine geographical heterogeneity. Such practice has been widely applied to studies examining neighborhood effects. Naturally, the variance of place effects across counties is already a lower bound for the total variance of neighborhood effect. Previous studies have revealed considerable variation in outcomes across counties [22]. Thus, focusing on county-level variations significantly increases the robustness and reduces the noise in estimation when discussing geographic variation. It is intuitive and evidenced that the selection effects account for a larger fraction of variance in permanent resident's COVID-19 infections at smaller geographics too since families and individual reaction to NPIs are more likely to sort geographically (e.g., social distance within the community or county area) within rather than across the regions or states.

Theoretically, Macintyre and colleagues [23] conceptualized the "place effects" on health into compositional (e.g., individual socioeconomic characteristics), contextual (e.g., healthcare structures and accessibility), and collective (e.g., cultural values and shared norms) dimensions. Macintyre's frame has significant potential to systematically understand the geographical patterning of NPIs' impact across counties. One prior study based on county-equivalent areas in the USA found that county-level factors, including demographic distribution, socioeconomic status, health care access, influenced the cumulative COVID-19 confirmed cases and deaths [24]. However, it remained unclear whether county-level factors influence the NPIs implementation, reflecting through the post-NPIs COVID-19 growth rates trend. Such information is critical to guide future tailored policy interventions.

This study leveraged multiple public datasets and a machine learning approach to construct the county-level spatial-temporal prediction model of COVID-19 in NYS. To guide a more effective reopening plan while controlling the transmission, the current study aims at 1) identifying hotspots of the growth rate of COVID-19 incidence among the 57 counties and New York City (NYC) in NYS over time, and 2) examining the association of COVID-19 growth rates after eight critical NPIs time points and most relevant county-level sociological predictors. It is hypothesized that there will be disproportionate COVID-19 incidence in counties with higher poverty rates, larger percentages of racial/ethnic minority population, fewer health care resources, and more disadvantaged neighborhoods after implementing the eight NPIs. This study is among the first to investigate such associations across NYS at the county levels while accounting for geographical and temporal variations.

NYS was chosen for the current study because of its generalizability to other states and similar metropolitan areas outside of the US. Specifically, New York State has increasingly diverse populations demographically, geographically, and socioeconomically [25], which allows sufficient variance to detect possible geographic differences and neighborhood effects [13]. Besides, NYS quickly became an epicenter of the pandemic in the world since the outbreak in March, while the government also responded timely with multiple NPIs. This offers an excellent opportunity to detect the geospatial variations in COVID-19 incidence after NPIs and temporal changes from the initiation to peak and slow-down pandemic. The current study looks at NYS, hoping to provide some insights for the rest of the world.

Methods

Data Resources

All variables of interest in this study were extracted from government open-access databases and public datasets, including Health Data in NYS [26], US Health Data [27], US FBI Crime Data [28], and American Community Survey (ACS) [29]. These data were aggregated to the county level for further analysis.

Selected Time Points

Based on our research hypotheses and interest, eight critical policy reactions announced by the state government were determined, and the policy announcement dates were selected for data analysis. Dates of NPIs in this study, collected by tracking news coverage and verifying news reports against government policies, are consistent with previous studies on anti-contagion systems [30] and other NPIs [7, 31]. Since the reduction in incidence was found to occur within a lag given that the incubation period for COVID-19 was 2-14 days [32, 33], eight post-14-day lagged time points were also included (Appendix 1).

5-day Moving Average Growth Rate of COVID-19 Incidence

The daily growth rates of cumulative COVID-19 cases from the previous day were extracted by regions (57 counties and NYC) in NYS from the US Health Data through the Social Explorer platform [27]. To reduce the impact of “noise” and obtain a smoothed estimate of growth rate, 5-day moving average growth rate was calculated for each selected time point by averaging the values of the selected day, two days before, and two days after [34].

County-level Predictors

Guided by Macintyre’s framework of place effect on health [23], 36 predictors across compositional, contextual, and collective factors were included. *Compositional factors* refer to the sociodemographic characteristics of individuals living in a particular place [23]. In this study, 5-year estimated county-level factors, such as population density (per square mile), the proportion of male, individuals who were older than 45 years (%), Black (%), Asian (%), Latino (%), children living in single-mother family (%), population in group quarters (%), noncitizen (%), living alone (%) in occupied and rental houses, living alone (%) in the same occupied and rental houses, and population who used public transportation to work (%) were extracted from the 2014-2018 ACS and aggregated to county levels [29]. To assess area socioeconomic status (SES), 5-year estimated factors including poverty, low education, unemployment, and low working class were considered [29, 35]. *Contextual factors* are defined as the broader social, economic, and physical opportunities in a region [23]. Socioeconomic opportunities were operationalized by 5-year estimated rental houses, Gini index, vehicle access at occupied and rental houses, crowd occupied and rental houses, and house sizes at occupied and rental houses extracted from 2014-2018 ACS [29, 35]. According to the United States Department of Agriculture Economic Research Service, counties with less than 250 000 population were classified as rural areas while others were considered urban [36]. Physical and health opportunities were represented by the proportion of people not having health insurance (%, extracted from the ACS), rates or primary care provider, mental healthcare provider, limited access to a doctor due to costs, fair or poor health population, and year of potential life lost per 100,000 people (%, extracted from the 2016 US Health Data) [37]. *Collective factors* refer to the socio-cultural features of a region (e.g., safety and religion) [23]. County safety was measured by violent crime and property crime rates (extracted from the U.S. Crime Data, FBI) [28]. Religious environment was assessed by the percentage of religious adherents in 2010 by (obtained from the U.S. Religion Data) [28].

5-day Moving Average Growth Rate of COVID-19 Testing

COVID-19 testing data was accessed from the Health Data of NYS [26]. 5-day moving average growth rate of COVID-19 testing was calculated for each selected time point.

Statistical Analysis

Cumulative COVID-19 incident infection rates (i.e., 5-day moving average growth rate of COVID-19 incidence) and testing rates (i.e., 5-day moving average growth rate of COVID-19 testing rates) were linked with population-level and community-level sociological variables by unique county federal information processing standards codes. First, descriptive statistics for all the county-level factors by relevant time frames were reported using 5 quantile values (i.e., minimum, 25^{th} percentile, 50^{th} percentile, 75^{th} percentile, and maximum). Second, geospatial differences of COVID-19 growth rate at county levels for each selected time point were conducted. Temporal trends of COVID-19 growth rates in NYS were demonstrated by heatmap using the "ggplot2" package. Third, generalized mixed effect least absolute shrinkage and selection operator (LASSO) was conducted using the "glmmLasso" package to examine the impact of county-level factors on the same-day 14-day-lagged COVID-10 growth rates. Randoms effect was used to account for the repeated measure in the same county. Data were normalized before the analysis. Relevant COVID-19 testing growth rates on dates respective to the 16 selected time points were adjusted. To evaluate the model fit, the absolute mean square error between original and predicted values were reported. Tuning parameter was chosen using the smallest Akaike Information Criterion (AIC) value [38]. All analyses and plots were conducted using *R* version 3.6.3 (The *R* Foundation). Statistical significance was set at $p<0.05$, and all tests were two-tailed.

Results

Descriptive Statistics

Among the 57 counties and NYC, the growth rate of COVID-19 infection increased during March and achieved the peak by the end of March, followed by a decrease until June (table 2). The second and fourth quantiles of COVID-19 growth rate were 0% and 41.70% (interquartile range [IQR]=41.70%) on March 20th, 1.12% and 2.99% (IQR=1.87%) on April 20th, and 0.13% and 0.81% (IQR=0.68%) on June 22nd (table 2). Across the regions, nearly half of the population were male and/or equal to or older than 45 years (table 2). Large county variation exists in the population density (25th percentile: 69.70 per square mile; 75th percentile: 280.71 per square mile; IQR: 211.01 per square mile), mental health provider rate (25th percentile: 113.31 per 100 000 people; 75th percentile: 220.19 per 100 000 people; IQR: 106.89 per 100 000 people), years of potential life lost rate (25th percentile: 5675.00 per 100 000 people; 75th percentile: 6638.10 per 100 000 people; IQR: 963.10 per 100 000 people), and property crime rate (25th percentile: 155.73 per 100 000 people; 75th percentile: 734.46 per 100 000 people; IQR: 578.73 per 100 000 people).

[Insert Table 1 Here]

Spatial and Temporal Distribution of County-level COVID-19 Growth Rate

Spatial disparity and temporal trend of county-level COVID-19 growth rate were identified (Figure 1, Figure 2). At the beginning of the COVID-19 outbreak, the highest growth rates could be found in NYC and its neighboring counties (e.g., Westchester and Nassau). Then, the highest growth rates changed from these areas to the mid-Hudson region (e.g., Sullivan, Rockland, Orange, and Putnam) by the end of March. During April and May, the COVID-19 outbreak moved from the mid-Hudson region to the Central and Northeast regions. Counties in Central (e.g., Tioga, Madison, Oneida, Onondaga, and Cayuga) and Northeast (e.g., Hamilton, Herkimer, Fulton, Essex, and Washington) regions showed relatively high growth rates during this period. From May to June, the influence of COVID-19 infection expanded to

Western (e.g., Orleans, Chautauqua, Yates, Seneca, and Niagara) and Northern (e.g., St. Lawrence, Lewis, and Franklin) regions of NYS.

[Insert Figures 1 & 2 Here]

Associations between County-level Predictors and Post-NPI COVID-19 Growth Rates

Results of LASSO selected 16 of the 36 proposed predictors (Table 2). Among these predictors, generalized linear mixed effected regression revealed that proportions of Black (β=1.40 [0.35–2.45]) and Latino (β=1.74 [0.39–3.09]) population were positively associated with the same-day county-level COVID-19 growth rate. In contrast, counties with greater Gini index (β=-1.49 [-2.83–-0.15]), larger proportions of the population using public transportation to work (β=-1.57 [-2.85–-0.30]), higher rates of mental health provider (β=-0.94 [-1.64–-0.25]) were inclined to have lower same-day COVID-19 growth rate.

After incorporating the 14-day lagged time points (16 selected time points included in total) into the existing model for further analysis, 30 of 36 predictors were selected into the generalized linear mixed effect model by LASSO, and 6 variables could significantly predict county-level COVID-19 growth rate. Counties with a larger concentration of population aged 45 years and above (β=3.25[0.17–6.32]), living alone at residential houses (β=3.31 [0.39–-6.22]), and proportions of crowded residential houses (β=6.15 [2.15–10.14]) were more likely to have a higher growth rate of COVID-19 infection across the time. The proportion of people who were living alone at rental houses (β=-2.47 [-4.83–-0.12]) and percentage of mental health providers at local areas (β=-1.11 [-1.95–-0.28]) were negatively associated with the county-level COVID-19 growth rate during the study period. The final model fitted the data very well (Figure 3), given that mean of absolute differences between actual and predictive growth rate was 7.01% (ranging from 0.00% to 134.02%).

[Insert Table 2 Here]

[Insert Figure 3 Here]

Discussion

Using aggregated county-level data from multiple public datasets, findings of the current study revealed the shifts of hotspots of high COVID-19 growth rates in NYS from NYC (in March) to the Central-Northeast region (from April to May), and Western-Northern areas (May and June). Additionally, this study found high correlates between COVID-19 growth rates and county-level sociodemographic characteristics, including counties with less health care resources, crowded occupied houses, racial/ethnic minorities, economic disparities, group quarters, and counties relying on public transportation to work. This is the first study applying the sociological framework and time-varying analyses with a machine learning technique to identify influential predictors of COVID-19 growth rates. Findings contribute to informing the targets when planning the next steps of NPIs to reduce the spread of COVID-19 incidence. Methodologically speaking, the method of this study maximized the use of multiple geographical and temporal data sources and detected the most relevant county-level features correlated with COVID-19 growth rates, accounting for eight time points after implementing specific NPIs.

Findings of the current study identified spatial and temporal trends of COVID-19 growth rates, shifting from the concentration in the NYC area at the beginning of COVID-19 outbreak towards the mid-Hudson region around March, followed by moving to the Central and Northeast region between April and May, and moving to the Western and Northern areas between May and June. The changing hotspots could be a result of effective NPIs in controlling the increase of COVID-19 incidence. In particular, following the closure of schools (March 15, 2020), all non-essential businesses (March 20, 2020), parks (April 1, 2020), and subway services (April 20, 2020), large social gatherings were limited, and social distancing of at least six feet was mandatory. Previous empirical studies in the UK [5], China [2], Italy [39], and Brazil [40] indicated reducing physical contacts and increasing stringent social distancing as effective policies to control the virus transmission, growth rates, and mortality. A recent systematic review and meta-analysis also demonstrated lower transmission of viruses with one meter or more physical distancing (*n*=10 736, pooled adjusted odds ratio 0.18 [0.09−0.38]) among 172 observational studies across 16 countries [41]. The geospatial shift towards the Center and Western region of NYS in the later periods may be explained by the lower base rates of incidence at the beginning of our assessments, slower policy reactions targeting their healthcare resources and socioeconomic disparities, and lower fidelity of NPIs implementation [2, 42-44].

Mental health provider percentages were consistently and negatively related to the same-day and 14-day lagged growth rates of COVID-19. The availability of mental health providers reflects the magnitude of healthcare resources at county levels [45]. In particular, most mental health providers (e.g., psychologists) are partnered with primary care practices [45, 46]. Previous studies suggested greater rates of physicians and mental health providers may serve as a proxy for general quality of health care in a county (e.g., Westchester and NYC have more mental health providers than others), and hence the better capacity to flatten the curve of incidence [47]. Besides, the surging cases among these countries in the very early stage of the COVID-19 outbreak also triggered rapid policy responses that reduced the health burden.

Economic and racial/ethnic disparities in the growth rates of COVID-19 incidence were documented. Contrary to a prior finding using state-level Gini Index [48], the negative relationship between the Gini Index and COVID-19 growth rates at county levels revealed the need to investigate geographical variations in local resources. In particular, counties with high income inequality (e.g., NYC and Westchester) also received intensive NPIs and had a good reaction to these interventions, which were sufficient to bring the growth rate down and fulfilled the needs for medical devices, healthcare workers, and stringent lockdown strategies [49]. This study adds value to the existing literature by examining the county-level Gini index and providing information for policymakers to engage extra efforts to mitigate future COVID-19 on the most financially vulnerable counties.

The growth rates of incidence during the eight selected time points were higher in counties with higher percentages of Black (i.e., Westchester, NYC, Albany) and Latinx population (i.e., Westchester, NYC, Suffolk, Rockland). This is consistent with earlier studies that racial/ethnic minorities are more vulnerable to the COVID-19 pandemic [47]. The disproportionate impact of the COVID-19 epidemic on Black and Latinx populations has been structural factors preventing them from adhering to the social distancing policy. Most of them make up “essential workers” or “front-line workers,” including public transit employees, grocery workers, and custodial staff, increasing their likelihood of being infected [50]. Besides the social and structural determinants of health, racisms, discrimination, economic and educational disadvantages, health care access may also be the underlying cause of health disparities in COVID-19 [15]. For example, racial minorities are also more likely to be part of residentially segregated communities where health care resources are relatively inadequate [15]. The non-significant influences of Black and Latinx distribution after accounting for the 14-day lagged time points may reflect the effectiveness of early NPIs in flattening the increasing incidence.

The greater incidence growth rate of COVID-19 among counties with greater percentages of housing occupied by more than three residents is a unique result. While previous reports were stating the positive relationship between density and transmission rates in general [4], we investigated the differences between occupied housing and rental housing, which partially accounted for the social mobility after the COVID-19 outbreak. In particular, residents in occupied housing may be less flexible and mobile than those living in rental units. Hence, with low social mobility, it is more likely to result in greater speed of incidence after the first case tested positive in the household.

The positive influence of public transit service use on the COVID-19 growth rate after accounting for all the 16 time points was consistent with a previous study using geolocated cell phone data in King County, Washington [51]. Changes in travel behaviors using public transportation can partially explain the gap in SES between higher- versus lower-income workers. In particular, NYC ranked the highest among the rates of workers who commuted using public transportation. More than half of the essential workers in NYC are foreign-born, and most of them continued to have limited options for other means of transportation while commuting to work, even after the social distancing mandate was issued. Public health leaders need to address the vulnerabilities in accessing safe public transportation, especially in counties with dense populations and significant inequality.

The results further supported the fact that the impact of population density on COVID-19 incidence growth rate was more complex [47]. Among high-density counties like NYC, an early burst of infection cases was associated with a high level of NPIs that subsequently mitigate the epidemic. For example, the percentage of group quarters in the county has positive effect on the COVID-19 growth rate but this effect is not statistically significant. The lack of significance may be due to early policy reaction to enhance infection control measures in nursing homes in NYS, or the small variance in the percentage of group quarters from county to county [52]. Hence, it is essential to use a data-driven approach to predict future incidence growth to help policymakers making the right decisions in a timely way.

Public Health Implications

Targeting specific county-level factors and high-risk subpopulations with specific NPIs at different periods of COVID-19 pandemic is key to effectively prevent future transmission while managing the economic downturn. This study suggests that state policymakers shall invest more health resources in counties with low mental health providers, high economic inequality, large percentages of Black and Latinx populations, more crowded occupied housing, and more workers commuting to work by public transportation. Both short-term and long-term policies, locally and federally, are needed to focus on tackling barriers to health care services and ensuring immediate, reliable, affordable, and acceptable services to socioeconomically disadvantaged communities. Readily accessible testing, feasible work conditions, paid sick leave for workers feeling unwell, stable employment, appropriate access to personal protective equipment, and inclusion in clinical trials for vaccines and treatment in hospitals serving economically disadvantaged counties are critically in need now. Specifically, to relieve the burdens of frequent public transportation use and minimize social interactions, the government shall consider providing more diverse and social-distancing transportation options. For instance, in high-infection cities such as Wuhan, China, private shuttles were available for essential workers for work-related transportation [53]. To control infections during traveling in holiday seasons, strong government policies in closing public transportation services were found effective in controlling the infection of COVID-19 in China [54]. To reduce the potential danger of fast infection in crowded occupied housing, providing alternative and temporary housing for rental purposes could be an optional NPIs to reduce the burden of COVID-19 incidence. Additional housing policies, such as stopping enforcement of evictions due to COVID-19 related issues or freezing mortgage payments, were also found to reduce the growth rates of COVID-19 infections [7, 30, 55]. For frontline healthcare professionals in China, all food was delivered in their own hotel rooms at work [54].

One unique contribution of the current study is highlighting the needs for public health policymakers and legislators to address the structural and historical vulnerabilities associated with pre-existing socioeconomic and racial/ethnic disparities. For example, legislative policies targeting healthcare coverage, racisms, and all forms of discrimination and immigrant rights shall be considered while planning NPIs to address structural and historical health gaps. In addition, innovative technologies, including remote telehealth service delivery, are encouraged to reduce the barriers of accessing healthcare treatment and prevention services by geographic locations (e.g., urban versus rural contexts) and sociodemographic characteristics (e.g., White residents versus racial/ethnic minorities). For example, previous studies found online mental health services being used for the COVID-19 epidemic in China facilitated the development and improved the effectiveness of public emergency NPIs [56]. Finally, it is important to use geographic-specific monitoring systems and process evaluation to address the spatial and temporal trends of incidence when implementing NPIs [30].

Strengths and Limitations

There are several research values of the analytic methods of this study worth mentioning. First, while previous machine learning predictive models in public health applications have been criticized for being a-theoretical and hard to interpret results [57], this study associated sociological theory of the “place effects” [23] and included the examination of compositional, contextual, and collective indicators of COVID-19 incidence over time. Findings also provide opportunities for generating new insights into further policy and theory development.

Second, this study carefully calibrated COVID-19 incidence and testing rates by calculating the smoothed estimate of 5-day moving average growth rates. It is further advantageous by accounting for both same-day and 14-day-lagged incidence rates, consistent with previous findings that the effects of introducing and lifting NPIs delayed by 1–3 weeks [33]. Besides, since the total positive cases depend on the total number of testing conducted [58], this study further controlled the 5-day moving average growth rates of the COVID-19 testing over time. Such adjustment is needed to reflect demographic disparities in COVID-19 testing in the most vulnerable communities in New York [8, 59]. The current study also advanced the previous methods by augmenting county-level indicators from multiple publicly available and regularly collected administrative data, which not only allows for future replications with more extended periods of examination but opportunities to understand various dimensions of neighborhood effects. Previous studies also found that county-based covariates resulted in the best performing model (max R^2= 0.90) than State-based and COVID-based (max R^2= 0.85) clustering strategies [58].

Thirdly, this study considers both temporal and geographical variations in COVID-19 growth rates to address the heterogeneity of the counties regarding socioeconomic characteristics, which is valuable and sheds light on one importance of considering health inequalities. One previous study on machine learning-based predictions of the COVID-19 outbreak in China also showed that modeling for geo-specific similarities could improve predictive performance [60].

Lastly, LASSO is advantageous to conduct variable selection effectively. As LASSO provides spare weights by driving small weights to zero, this paper can find the strongest predictors of COVID-19 growth rates [61-64]. LASSO also shows good performance to handle multicollinearity [62, 65] and gains better prediction accuracy because the penalization term decreases the model's over-fitting [38, 63, 66]. LASSO adds an $\ell 1$ regularization term to select only the most important features to include in the model. This fits the aims of this study to identify policy, sociodemographic, and geographic factors that are closely related to the growth rate of COVID-19 while not aiming at building predictive models. In particular, the generalized mixed effect LASSO model outperformed other machine-learning methods (e.g., classification trees or support vector regression) when dealing with the nature of study data (i.e., longitudinal, clustered in the county level) [38]. This study further applied nested cross-validation to choose the LASSO penalty as recommended to improve validation [38, 67]. The method has been widely used by studies with similar study objectives [68-70].

This study has limitations. First, COVID-19 incidence rates were calculated depending on the date of official issues of key NPIs. While both same-day and post-14-day rates were accounted for, the actual effects of the exact policy may not be accurately determined. Besides, the incidence rates may fluctuate between the announcement dates and actual implementation dates. The final choices of the NPI dates were consistent with government and news reports, previous studies [7, 30, 55, 71] and publicly available COVID-19 US state policy database [31]. Second, the NYS government implemented multiple NPIs in a short timeframe to control the outbreak, while the selected time did not account for compound NPIs. Additionally, the fact that some sociodemographic factors were not significantly associated with incidence growth in the results may result from intensive NPIs in the early stage of the outbreak, which substantially mitigated the growth rates [30]. Third, other county-level factors, such as the nature of specific employment (e.g., full-time, part-time, self-employed) or business categories (e.g., retail, wholesale, restaurant) in the county, were not included to avoid over-complicated models. Lastly, although this is the first study to have comprehensive and theory-based county-level factors from multiple datasets (e.g., 2018 ACS, 2016 U.S. Health Data, and 2020 U.S. Religion Census), some factors were collected before the onset of the COVID-19 pandemic. It is possible that these factors changed slightly between the survey collection dates and the current status.

Conclusion

The emergence of COVID-19 has caused evolving global public health and economic crisis. Without reliable and effective pharmaceutical agents, leading public health, medical and political community rely on NPIs to reduce the burden of rapidly increasing transmission. This study, based on the case of NYS – the epicenter of COVID-19 cases worldwide, advanced the current knowledge of growth rates of COVID-19 after implementing NPIs by quantifying temporal and spatial variations in response to NPIs across counties. The present study also uniquely identified county-level factors associated with the average growth of COVID-19 incidence rates. Findings highlight the need to consider structural and historical factors to tackle the epidemic with efficacy provides a valuable perspective, which is too often forgotten in the medical literature. The current study further suggests the public health importance of tailoring NPIs towards geographical characteristics to improve the effectiveness in reducing future outbreak.

Acknowledgement

Role of the Funding Source

The funding source had no role in study design, data collection, data analysis, interpretation, and manuscript drafting. The corresponding author had full access to all the data used in the study and had full responsibility for the decision to submit for publication.

Declarations of Interests:

I declare no competing interests.

Informed Consent:

Informed consent was obtained from all individual participants included in the study.

References

1. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. Lancet Infect Dis. 2020;20(5):533-4.
2. Pan A, Liu L, Wang CL, Guo H, Hao XJ, Wang Q, et al. Association of Public Health Interventions With the Epidemiology of the COVID-19 Outbreak in Wuhan, China. JAMA. 2020 May;323(19):1915-23. PMID: WOS:000536934300018. doi: 10.1001/jama.2020.6130.
3. Leung K, Wu JT, Liu D, Leung GM. First-wave COVID-19 transmissibility and severity in China outside Hubei after control measures, and second-wave scenario planning: a modelling impact assessment. Lancet. 2020.
4. Prem K, Liu Y, Russell TW, Kucharski AJ, Eggo RM, Davies N, et al. The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: a modelling study. Lancet Public Health. 2020.
5. Davies NG, Kucharski AJ, Eggo RM, Gimma A, Edmunds WJ. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. Lancet Public Health. 2020 Jul;5(7):e375-e85. PMID: 32502389. doi: 10.1016/s2468-2667(20)30133-x.
6. Giordano G, Blanchini F, Bruno R, Colaneri P, Di Filippo A, Di Matteo A, et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. Nature Medicine. 2020:1-6.
7. Jay J, Bor J, Nsoesie EO, Lipson SK, Jones DK, Galea S, et al. Neighbourhood income and physical distancing during the COVID-19 pandemic in the United States. Nature Human Behaviour. 2020 2020/11/03. doi: 10.1038/s41562-020-00998-2.
8. Borjas GJ. Demographic determinants of testing incidence and COVID-19 infections in New York City neighborhoods. National Bureau of Economic Research, 2020 0898-2937.
9. Whittle RS, Diaz-Artiles A. An ecological study of socioeconomic predictors in detection of COVID-19 cases across neighborhoods in New York City. BMC Medicine. 2020 2020/09/04;18(1):271. doi: 10.1186/s12916-020-01731-6.
10. Sampson RJ, Morenoff JD, Gannon-Rowley T. Assessing “neighborhood effects”: Social processes and new directions in research. Annual review of sociology. 2002;28(1):443-78.
11. Braveman PA, Egerter SA, Cubbin C, Marchi KS. An approach to studying social disparities in health and health care. American Journal of Public Health. 2004;94(12):2139-48.
12. Osypuk TL, Acevedo-Garcia D. Beyond individual neighborhoods: A geography of opportunity perspective for understanding racial/ethnic health disparities. Health & Place. 2010 2010/11/01/;16(6):1113-23. doi: https://doi.org/10.1016/j.healthplace.2010.07.002.

13. Arcaya MC, Tucker-Seeley RD, Kim R, Schnake-Mahl A, So M, Subramanian S. Research on neighborhood effects on health in the United States: a systematic review of study characteristics. Social Science & Medicine. 2016;168:16-29.

14. Hatef E, Chang H-Y, Kitchen C, Weiner JP, Kharrazi H. Assessing the Impact of Neighborhood Socioeconomic Characteristics on COVID-19 Prevalence Across Seven States in the United States. Front Public Health. 2020;8:571808-. PMID: 33072710. doi: 10.3389/fpubh.2020.571808.

15. Hooper MW, Nápoles AM, Pérez-Stable EJ. COVID-19 and racial/ethnic disparities. Jama. 2020.

16. Benitez J, Courtemanche C, Yelowitz A. Racial and ethnic disparities in Covid-19: Evidence from six large cities. Journal of Economics, Race, and Policy. 2020:1-19.

17. Jay J, Bor J, Nsoesie E, Lipson SK, Jones DK, Galea S, et al. Neighborhood income and physical distancing during the COVID-19 pandemic in the US. medRxiv. 2020.

18. Hilmers A, Hilmers DC, Dave J. Neighborhood disparities in access to healthy foods and their effects on environmental justice. American journal of public health. 2012;102(9):1644-54.

19. Hong B, Bonczak B, Gupta A, Thorpe L, Kontokosta CE. Exposure Density and Neighborhood Disparities in COVID-19 Infection Risk: Using Large-scale Geolocation Data to Understand Burdens on Vulnerable Communities. arXiv preprint arXiv:200801650. 2020.

20. Adhikari S, Pantaleo NP, Feldman JM, Ogedegbe O, Thorpe L, Troxel AB. Assessment of community-level disparities in coronavirus disease 2019 (COVID-19) infections and deaths in large US metropolitan areas. JAMA network open. 2020;3(7):e2016938-e.

21. Bowleg L. We're Not All in This Together: On COVID-19, Intersectionality, and Structural Inequality. American Journal of Public Health. 2020 2020/07/01;110(7):917-. doi: 10.2105/AJPH.2020.305766.

22. Chetty R, Hendren N. The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects. The Quarterly Journal of Economics. 2018;133(3):1107-62. doi: 10.1093/qje/qjy007.

23. Macintyre S, Ellaway A, Cummins S. Place effects on health: how can we conceptualise, operationalise and measure them? Social science & medicine. 2002;55(1):125-39.

24. Killeen BD, Wu JY, Shah K, Zapaishchykova A, Nikutta P, Tamhane A, et al. A County-level Dataset for Informing the United States' Response to COVID-19. arXiv preprint arXiv:200400756. 2020.

25. Manganello JA, Gerstner G, Pergolino K, Graham Y, Strogatz D. Understanding digital technology access and use among New York state residents to enhance dissemination of health information. JMIR public health and surveillance. 2016;2(1):e9.

26. New York State Statewide COVID-19 Testing, Health Data in New York State. 2020.
27. Post C, Byron K. Women on Boards and Firm Financial Performance: A Meta-Analysis. Acad Manage J. 2015 2015/10;58(5):1546-71. doi: 10.5465/amj.2013.0319.
28. Ehsan AM, De Silva MJ. Social capital and common mental disorder: a systematic review. J Epidemiol Community Health. 2015 Oct;69(10):1021-8. PMID: 26179447. doi: 10.1136/jech-2015-205868.
29. American Community Survey 2014-2018 (5-Year Estimate). 2018.
30. Hsiang S, Allen D, Annan-Phan S, Bell K, Bolliger I, Chong T, et al. The effect of large-scale anti-contagion policies on the COVID-19 pandemic. Nature. 2020 2020/08/01;584(7820):262-7. doi: 10.1038/s41586-020-2404-8.
31. Raifman J, Nocka K, Jones D, Bor J, Lipson S, Jay J, et al. COVID-19 US state policy database. 2020.
32. National Center for Immunization and Respiratory Diseases (NCIRD) DoVD. Symptoms of Coronavirus. 2020.
33. Li Y, Campbell H, Kulkarni D, Harpur A, Nundy M, Wang X, et al. The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries. The Lancet Infectious Diseases. 2020.
34. Johns Hopkins University CRC. Have countries flattened the curve? 2020.
35. Niyonsenga T, Trepka MJ, Lieb S, Maddox LM. Measuring socioeconomic inequality in the incidence of AIDS: rural-urban considerations. AIDS Behav. 2013 Feb;17(2):700-9. PMID: 22711226. doi: 10.1007/s10461-012-0236-8.
36. Service USDoAER. Rural-Urban Continuum Codes. 2013.
37. Taylor M, Perakakis P, Trachana V, Gialis S. Rankings are the sorcerer’s new apprentice. Ethics Sci Environ Polit. 2014 2014/05/15;13(2):73-99. doi: 10.3354/esep00146.
38. Schelldorfer J, Meier L, Bühlmann P. Glmmlasso: an algorithm for high-dimensional generalized linear mixed models using ℓ1-penalization. Journal of Computational and Graphical Statistics. 2014;23(2):460-77.
39. Chintalapudi N, Battineni G, Sagaro GG, Amenta F. COVID-19 outbreak reproduction number estimations and forecasting in Marche, Italy. Int J Infect Dis. 2020 May 11;96:327-33. PMID: 32437930. doi: 10.1016/j.ijid.2020.05.029.
40. Valenti VE, de Lemos Menezes P, de Abreu ACG, Vieira GNA, Garner DM. Social distancing measures may have reduced the estimated deaths related to Covid-19 in Brazil. Journal of Human Growth and Development. 2020;30(2):164-9.
41. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. Lancet. 2020;395(10242):1973-87. doi: 10.1016/S0140-6736(20)31142-9.

42. Richardson S, Hirsch JS, Narasimhan M, Crawford JM, McGinn T, Davidson KW, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area. Jama. 2020.
43. Mollalo A, Vahedi B, Rivera KM. GIS-based spatial modeling of COVID-19 incidence rate in the continental United States. Science of the Total Environment. 2020 Aug;728. PMID: WOS:000541009200021. doi: 10.1016/j.scitotenv.2020.138884.
44. Wu JY, Killeen BD, Nikutta P, Thies M, Zapaishchykova A, Chakraborty S, et al. Changes in Reproductive Rate of SARS-CoV-2 Due to Non-pharmaceutical Interventions in 1,417 US Counties. medRxiv. 2020.
45. Kakuma R, Minas H, Van Ginneken N, Dal Poz MR, Desiraju K, Morris JE, et al. Human resources for mental health care: current situation and strategies for action. Lancet. 2011;378(9803):1654-63.
46. Saraceno B, van Ommeren M, Batniji R, Cohen A, Gureje O, Mahoney J, et al. Barriers to improvement of mental health services in low-income and middle-income countries. Lancet. 2007;370(9593):1164-74.
47. Hamidi S, Ewing R, Sabouri S. Longitudinal analyses of the relationship between development density and the COVID-19 morbidity and mortality rates: Early evidence from 1,165 metropolitan counties in the United States. Health & Place. 2020:102378.
48. Oronce CIA, Scannell CA, Kawachi I, Tsugawa Y. Association Between State-Level Income Inequality and COVID-19 Cases and Mortality in the USA. Journal of General Internal Medicine. 2020:1-3.
49. Bor J, Cohen GH, Galea S. Population health in an era of rising income inequality: USA, 1980–2015. Lancet. 2017;389(10077):1475-90.
50. van Dorn A, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. Lancet. 2020;395(10232):1243.
51. Brough R, Freedman M, Phillips D. Understanding Socioeconomic Disparities in Travel Behavior during the COVID-19 Pandemic. University of California, Irvine Department of Economics Working Paper Series. 2020.
52. Chowell G, Mizumoto K. The COVID-19 pandemic in the USA: what might we expect? Lancet. 2020;395(10230):1093-4.
53. Liu M, Cheng S-Z, Xu K-W, Yang Y, Zhu Q-T, Zhang H, et al. Use of personal protective equipment against coronavirus disease 2019 by healthcare professionals in Wuhan, China: cross sectional study. BMJ. 2020;369:m2195. doi: 10.1136/bmj.m2195.
54. Chen S, Yang J, Yang W, Wang C, Bärnighausen T. COVID-19 control in China during mass population movements at New Year. The Lancet. 2020;395(10226):764-6.
55. Courtemanche C, Garuccio J, Le A, Pinkston J, Yelowitz A. Strong Social Distancing Measures In The United States

Reduced The COVID-19 Growth Rate: Study evaluates the impact of social distancing measures on the growth rate of confirmed COVID-19 cases across the United States. Health Affairs. 2020:10.1377/hlthaff. 2020.00608.

56. Liu S, Yang L, Zhang C, Xiang Y-T, Liu Z, Hu S, et al. Online mental health services in China during the COVID-19 outbreak. The Lancet Psychiatry. 2020 2020/04/01/;7(4):e17-e8. doi: https://doi.org/10.1016/S2215-0366(20)30077-8.
57. Poggio T, Banburski A, Liao Q. Theoretical issues in deep networks. Proceedings of the National Academy of Sciences. 2020:201907369. doi: 10.1073/pnas.1907369117.
58. Kefayati S, Huang H, Chakraborty P, Roberts F, Gopalakrishnan V, Srinivasan R, et al. On machine learning-based short-term adjustment of epidemiological projections of covid-19 in US. medRxiv. 2020.
59. Lieberman-Cribbin W, Tuminello S, Flores RM, Taioli E. Disparities in COVID-19 Testing and Positivity in New York City. American Journal of Preventive Medicine. 2020 2020/09/01/;59(3):326-32. doi: https://doi.org/10.1016/j.amepre.2020.06.005.
60. Liu D, Clemente L, Poirier C, Ding X, Chinazzi M, Davis J, et al. Real-Time Forecasting of the COVID-19 Outbreak in Chinese Provinces: Machine Learning Approach Using Novel Digital Data and Estimates From Mechanistic Models. J Med Internet Res. 2020;22(8):e20285. PMID: 32730217. doi: 10.2196/20285.
61. Hafner CM. The Spread of the Covid-19 Pandemic in Time and Space. Int J Environ Res Public Health. 2020 May 28;17(11). PMID: 32481629. doi: 10.3390/ijerph17113827.
62. Januaviani TMA, Gusriani N, Joebaedi K, Supian S, Subiyanto S. The best model of LASSO with the LARS (least angle regression and shrinkage) algorithm using Mallow's Cp. World Scientific News. 2019;116:245-52.
63. McNeish DM. Using Lasso for Predictor Selection and to Assuage Overfitting: A Method Long Overlooked in Behavioral Sciences. Multivariate Behav Res. 2015;50(5):471-84. PMID: 26610247. doi: 10.1080/00273171.2015.1036965.
64. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society: Series B (Methodological). 1996;58(1):267-88.
65. Chong I-G, Jun C-H. Performance of some variable selection methods when multicollinearity is present. Chemometrics and intelligent laboratory systems. 2005;78(1-2):103-12.
66. Van de Geer SA. High-dimensional generalized linear models and the lasso. The Annals of Statistics. 2008;36(2):614-45.
67. Fan Y, Tang CY. Tuning parameter selection in high dimensional penalized likelihood. Journal of the Royal Statistical Society: SERIES B: Statistical Methodology. 2013:531-52.

68. Bremer V, Becker D, Kolovos S, Funk B, van Breda W, Hoogendoorn M, et al. Predicting Therapy Success and Costs for Personalized Treatment Recommendations Using Baseline Characteristics: Data-Driven Analysis. Journal of medical Internet research. 2018;20(8):e10275-e. PMID: 30131318. doi: 10.2196/10275.
69. Palmius N, Saunders KEA, Carr O, Geddes JR, Goodwin GM, De Vos M. Group-Personalized Regression Models for Predicting Mental Health Scores From Objective Mobile Phone Data Streams: Observational Study. Journal of medical Internet research. 2018;20(10):e10194-e. PMID: 30348626. doi: 10.2196/10194.
70. Cepeda MS, Reps J, Kern DM, Stang P. Medical Conditions Predictive of Self-Reported Poor Health: Retrospective Cohort Study. JMIR public health and surveillance. 2020;6(1):e13018-e. PMID: 31913130. doi: 10.2196/13018.
71. Lai S, Ruktanonchai NW, Zhou L, Prosper O, Luo W, Floyd JR, et al. Effect of non-pharmaceutical interventions to contain COVID-19 in China. Nature. 2020 May 4. PMID: 32365354. doi: 10.1038/s41586-020-2293-x.

Table 1. Descriptive statistics of coronavirus disease (COVID-19) 5-day moving average growth rates of infection, county-level factors, and 5-day moving average growth rates of testing in New York State based on the US Health Data through Social Explorer between March 7, 2020 and June 22, 2020.

Variables	Q1	Q2	Q3	Q4	Q5	IQR
5-day moving average growth rates of infection						
March 7	0	0	0	0	4	4
March 16	0	0	0	20	107.45	107.45
March 20	0	0	22.50	41.70	156.67	156.67
March 21	0	0	30	56.67	165.42	165.42
March 30	0	12.07	17.52	26.41	97.33	97.33
April 1	0	10.53	14.29	18.25	60	60
April 3	0	9.90	13.74	17.48	70	70
April 4	0	8.06	11.92	16.37	50	50
April 15	0	3.08	4.00	5.21	22.86	22.86
April 18	0	1.40	2.66	4.18	17.82	17.82
April 20	0	1.12	2.25	2.99	14.33	14.33
May 4	0	2.97	4.55	6.64	23.11	23.11
May 9	0	0.66	1.32	2.52	13.33	13.33
June 8	0	0.82	1.55	2.17	4.83	4.83
June 22	0	0.13	0.47	0.81	2.24	2.24
Compositional factors						
Population density (%)	2.66	69.70	108.72	280.71	28110.10	28107.44
Male (%)	47.73	49.09	49.69	50.34	54.87	7.14
Age (%)	34.44	44.40	46.53	48.42	62.93	28.49
Black (%)	0.60	1.58	3.28	5.97	19.72	19.12
Asian (%)	0.02	0.71	1.10	2.87	12.18	12.16
Latino (%)	1.55	2.66	3.55	6.91	29.48	27.93
Single-mother family (%)	7.21	9.73	10.83	11.74	17.93	10.72
Group quarter (%)	1.12	2.12	4.14	6.52	13.32	12.20
Low education (%)	5.77	8.92	10.11	12.12	17.86	12.09
Unemployment (%)	3.50	4.79	5.71	6.41	8.03	4.53
Poverty (%)	5.42	10.69	13.48	15.66	23.99	18.57

Public transportation to work (%)	0.25	0.73	1.20	3.13	53.24	52.99
Noncitizen (%)	0.31	0.93	1.67	3.43	15.02	14.71
Low working class (%)	33.85	41.73	46.98	49.62	54.77	20.92
No health insurance (%)	3.30	4.21	5.05	6.04	18.82	15.52
Residential alone (%)	19.92	27.25	29.42	30.77	35.65	15.73
Rental alone (%)	29.27	38.9	43.15	45.11	54.47	25.20
Residential same house within one year (%)	74.97	86.08	88.21	90.04	93.83	18.86
Rental same house within one year (%)	13.22	16.91	19.72	22.92	53.21	39.99
Contextual factors						
Rental houses (%)	14.32	25.03	28.06	32.27	62.38	48.06
Gini index	0.39	0.42	0.44	0.46	0.54	0.15
Residential car access (%)	4.27	7.29	8.97	11.17	48.96	44.69
Rental car access (%)	3.11	19.21	22.52	26.62	62.17	59.06
Residential crowd (%)	18.78	31.62	33.95	35.71	51.61	32.83
Rental crowd (%)	21.32	27.15	30.71	33.60	47.88	26.56
Residential house size	2.26	2.34	2.39	2.54	3.95	1.69
Rental house size	1.95	2.12	2.17	2.32	4.34	2.39
Primary care provider rate per 100 000 people	9.47	43.04	54.72	76.83	146.58	137.11
Mental health provide rate per 100 000 people	21.21	113.31	151.62	220.19	334.56	313.35
Limited access to doctor due to costs ($)	6944.47	7970.79	8393.14	8852.82	10354.75	3410.28
Fair or poor health (%)	9.80	12.1	12.65	13.10	18.72	8.92
Year of potential life lost	14.32	25.03	28.06	32.27	62.38	48.06
Collective factors						
Religious adherents (%)	15.66	34.53	39.70	46.35	72.11	56.45
Violent crime rate per 100 000 people	0.64	16.08	38.53	83.46	526.20	525.56
Property crime rate per 100 000 people	7.37	155.73	381.42	734.46	1989.60	1982.23
Rurality (%)						
5-day moving average growth rates of testing						
March 7	0	0	0	0.25	5.03	5.03
March 16	0.19	0.36	0.51	0.74	4.00	3.81
March 20	0.18	0.35	0.47	0.59	1.73	1.55

March 21	0.06	0.27	0.36	0.47	0.94	0.88
March 30	0.02	0.08	0.10	0.12	0.21	0.19
April 1	0.02	0.06	0.10	0.12	0.21	0.19
April 3	0.02	0.07	0.09	0.11	0.20	0.18
April 4	0.00	0.06	0.08	0.10	0.20	0.20
April 15	0.02	0.03	0.04	0.05	0.08	0.06
April 18	0.01	0.02	0.03	0.04	0.08	0.07
April 20	0.01	0.02	0.03	0.03	0.13	0.12
May 4	0.02	0.04	0.06	0.08	0.15	0.12
May 9	0.02	0.03	0.04	0.05	0.10	0.08
June 8	0.02	0.03	0.04	0.05	0.12	0.11
June 22	0.01	0.03	0.03	0.04	0.07	0.05

Notes: Q1=Minimum; Q2=25th percentile; Q3=50th percentile; Q4=75th percentile; Q5=Maximum; IQR=Interquartile range.

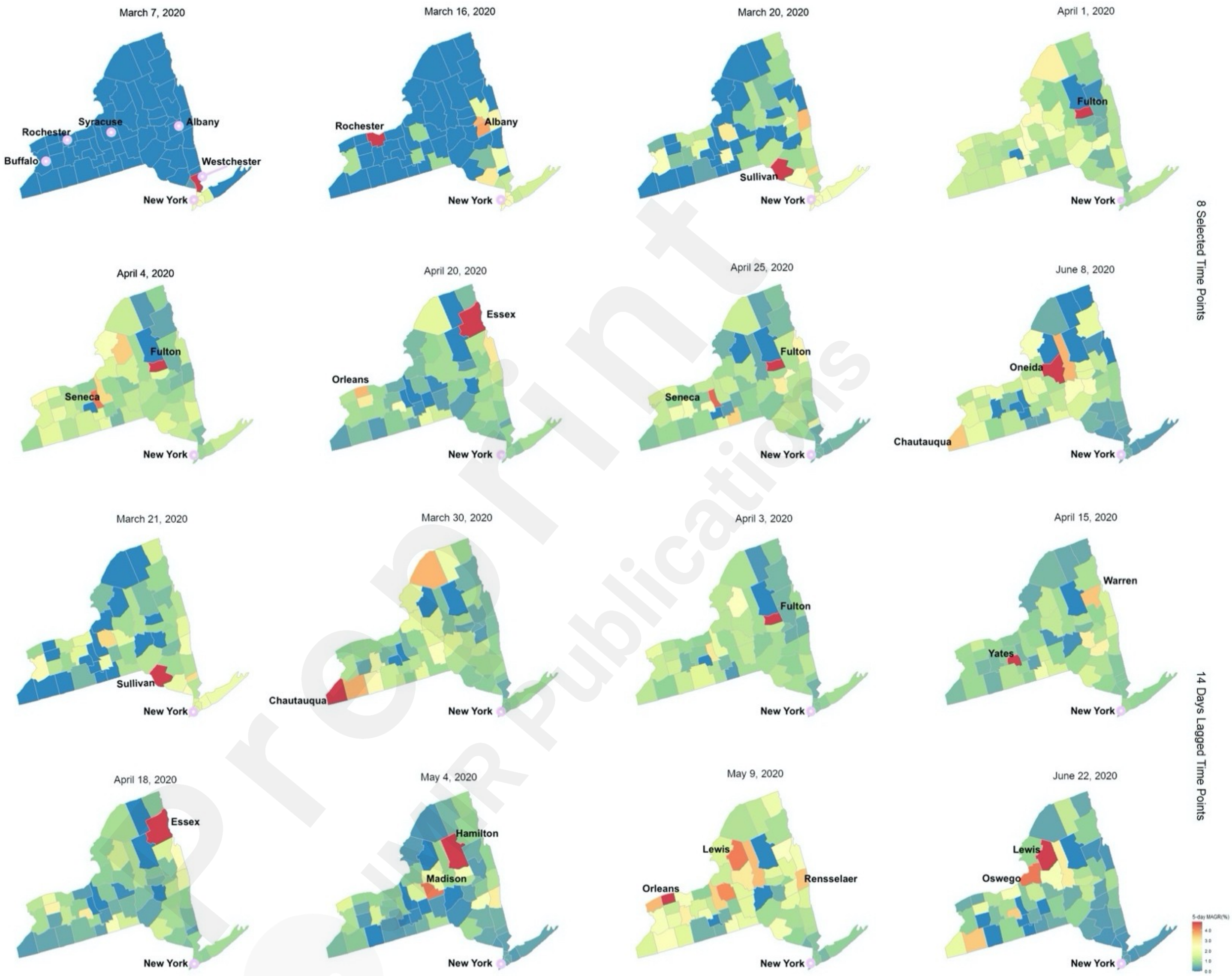
Table 2. Associations between county-level factors and 5-day moving average growth rates of coronavirus disease (COVID-19) infection rates on the same date of eight selected NPIs and 14-day-lagged dates of NPIs in New York State based on the US Health Data through Social Explorer between March 7, 2020 and June 22, 2020.

	Same date of NPIs			Same date and 14-day-lagged date of NPIs		
Predictors	*β*	*95% CI Lower*	*95% CI Upper*	*β*	*95% CI Lower*	*95% CI Upper*
Compositional factors						
Population density (%)	-1.67	-4.40	1.06	-1.65	-4.43	1.14
Male (%)	0	NA	NA	-0.25	-1.30	0.81
Age (%)	0	NA	NA	**3.25***	**0.17**	**6.32**
Black (%)	**1.40****	**0.35**	**2.45**	1.03	-0.25	2.30
Asian (%)	1.23	-0.40	2.86	-0.62	-2.81	1.57
Latino (%)	**1.74***	**0.39**	**3.09**	-0.30	-1.95	1.35
Single-mother family (%)	1.00	-0.17	2.17	0.06	-1.55	1.68
Group quarter (%)	0	NA	NA	0	NA	NA
Low education (%)	0	NA	NA	-0.60	-2.35	1.16
Unemployment (%)	0	NA	NA	0	NA	NA
Poverty (%)	0	NA	NA	0	NA	NA
Public transportation to work (%)	**-1.57***	**-2.85**	**-0.30**	0.38	-1.37	2.12
Noncitizen (%)	1.17	-1.27	3.61	2.02	-1.22	5.26
Low working class (%)	-0.81	-1.77	0.16	-0.46	-1.70	0.78
No health insurance (%)	0	NA	NA	-1.27	-2.87	0.34
Residential alone (%)	0	NA	NA	**3.31***	**0.39**	**6.22**
Rental alone (%)	0	NA	NA	**-2.47***	**-4.83**	**-0.12**
Residential same house within one year (%)	0	NA	NA	0	NA	NA
Rental same house within one year (%)	-0.09	-3.10	2.92	-0.08	-0.43	0.26
Contextual factors						
Rental houses (%)	0.34	-2.04	2.72	4.05	-1.01	9.10
Gini index	**-1.49***	**-2.83**	**-0.15**	-0.75	-2.57	1.07
Residential car access (%)	-1.81	-5.31	1.69	0	NA	NA
Rental car access (%)	0	NA	NA	0.84	-1.70	3.38
Residential crowd (%)	-1.56	-3.14	0.02	**6.15****	**2.15**	**10.14**

Rental crowd (%)	0	NA	NA	**-1.80****	**-3.75**	**0.14**
Residential house size	0	NA	NA	0.98	-1.40	3.36
Rental house size	0	NA	NA	0	NA	NA
Primary care provider rate per 100,000 people	0	NA	NA	-0.29	-1.46	0.88
Mental health provide rate per 100,000 people	**-0.94****	**-1.64**	**-0.25**	**-1.11****	**-1.95**	**-0.28**
Limited access to doctor due to costs ($)	0	NA	NA	0.09	-2.70	2.87
Fair or poor health (%)	0	NA	NA	-4.43	-10.14	1.28
Year of potential life lost	0	NA	NA	-0.18	-1.80	1.45
Collective factors						
Religious adherents (%)	-0.35	-1.19	0.48	-0.96	-2.19	0.27
Violent crime rate per 100,000 people	0	NA	NA	1.28	-1.19	3.74
Property crime rate per 100,000 people	0	NA	NA	0.03	-1.68	1.74
Rurality (%)	0.13	-0.13	0.39	-0.22	-1.03	0.60

Notes: Results from generalized linear mixed effect models with the LASSO penalty. NA=variables were not selected by generalized linear mixed effect model with LASSO penalty; 95% CI=95% confidence interval. *: p<.05; **: p<.01. NPIs= Non-pharmaceutical interventions. Bolded texts indicate statistically significant associations at p<.05.

Figure 1. Geographic distribution of coronavirus disease (COVID-19) infection rates on the same date of eight selected NPIs and 14-day-lagged dates of NPIs in New York State based on the US Health Data through Social Explorer between March 7, 2020 and June 22, 2020.



Note: NPIs=Non-pharmaceutical interventions. Eight time points for selected NPIs and 14-day lagged time points after the eight selected NPIs time.

Figure 2. Temporal distribution of coronavirus disease (COVID-19) infection rates on the same date of eight selected NPIs and 14-day-lagged dates of NPIs in New York State based on the US Health Data through Social Explorer between March 7, 2020 and June 22, 2020.

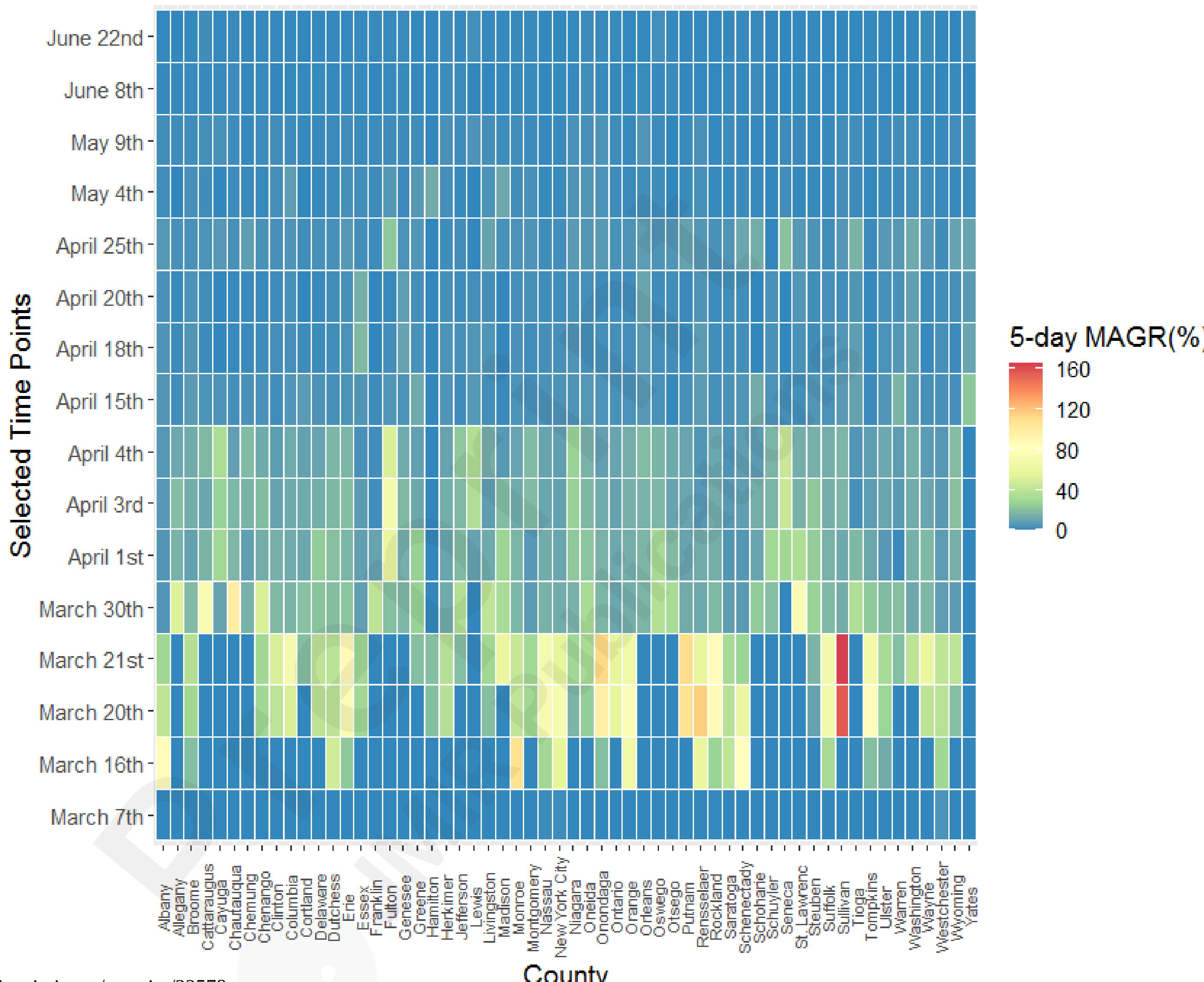
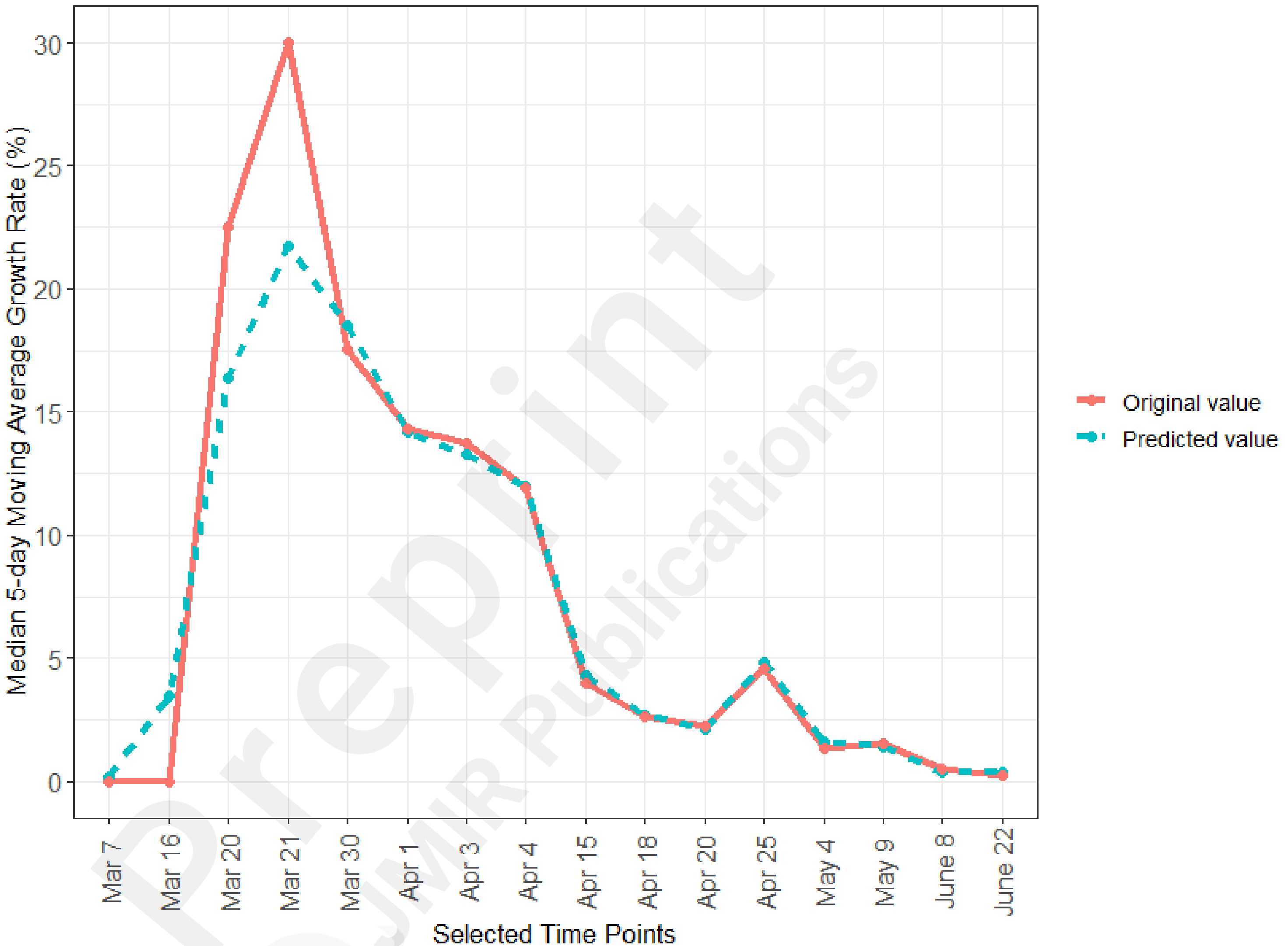


Figure 3 Predicted and original coronavirus disease (COVID-19) infection rates on the same date of eight selected NPIs and 14-day-lagged dates of NPIs in New York State based on the US Health Data through Social Explorer between March 7, 2020 and June 22, 2020.



Appendix 1 Dates and 14-day lagged dates of eight selected non-pharmaceutical interventions (NPIs) for controlling coronavirus disease (COVID-19) infection rates in New York State and intervention types between March 7, 2020 and June 22, 2020.

Time Points	Lag 14-day	New York State Government Policy	NPIs
March 7, 2020	March 21, 2020	State of Emergency Declare	*cordons sanitaire*
March 16, 2020	March 30, 2020	All school closed statewide	traffic restriction
March 20, 2020	April 3, 2020	New York State on PAUSE	social distancing home confinement
April 1, 2020	April 15, 2020	All parks in the city shut down	traffic restriction
April 4, 2020	April 18, 2020	Chinese government donate ventilator; USNS ship arrival	centralized quarantine
April 20, 2020	May 4, 2020	Subway service temporarily suspended	travel restriction
April 25, 2020	May 9, 2020	Testing expanded	testing, treatment, universal survey
June 8, 2020	June 22, 2020	Phase I Reopening	relaxing NPI

Note. Dates of time points were collected by tracking news coverage and verifying news reports against government websites.