



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.



## Original article

# Trajectory Modeling of Spatio-Temporal Trends in COVID-19 Incidence in Flint and Genesee County, Michigan



Thomas Walter Wojciechowski, PhD, Richard Casey Sadler, PhD, MPH\*, Zachary Buchalski, MS, Alan Harris, BSc, Danielle Lederer, MS, C Debra Furr-Holden, PhD

Michigan State University, Flint, MI

## ARTICLE INFO

## Article history:

Received 2 June 2021

Revised 29 November 2021

Accepted 5 December 2021

Available online 16 December 2021

## Keywords:

Covid-19

Gis

Health inequalities

Epidemiological methods

## ABSTRACT

**Purpose:** The establishment of community-academic partnerships to digest data and create actionable policy and advocacy steps is of continuing importance. In this paper, we document COVID-19 racial and geographic disparities uncovered via a collaboration between a local health department and university research center.

**Methods:** We leverage individual level data for all COVID-19 cases aggregated to the census block group level, where group-based trajectory modeling was employed to identify latent patterns of change and continuity in COVID-19 diagnoses.

**Results:** Linking with socioeconomic data from the census, we identified the types of communities most heavily affected by each of Michigan's two waves (in spring and fall of 2020). This includes a geographic and racial gap in COVID-19 cases during the first wave, which is largely eliminated during the second wave.

**Conclusions:** Our work has been extremely valuable for community partners, informing community-level response toward testing, treatment, and vaccination. In particular, identifying and conducting advocacy on the sizeable racial disparity in COVID-19 cases during the first wave in spring 2020 helped our community nearly eliminate disparities throughout the second wave in fall 2020.

© 2021 Elsevier Inc. All rights reserved.

## Introduction

The COVID-19 pandemic is a global crisis that is likely to impact public health for years to come [1,2]. The pervasive nature of the crisis and the novelty of the virus necessitate identification of drivers of local case spikes. Extensive work has already employed geographic information systems to address the dynamic spatial nature of COVID-19 spread [3]. Doing so locally can be extremely valuable in providing an understanding of why surges in cases occur when and where they do, as well as community characteristics of hotspots. This in turn allows for more accurate targeting of public health interventions and policy at the community level that may engender harm reduction.

The present study sought to identify factors that drove COVID-19 case surges during different points in the first year of the pandemic in Genesee County, Michigan. Genesee County is home to Flint, the epicenter of a water crisis that began in 2014 stemming from a state-directed infrastructure change that led to widespread lead-in-water and Legionella contamination [4–6]. Owing to this environmental injustice on top of years of disinvestment and discrimination that have created huge health disparities [7], this area has become increasingly of interest to public health researchers. At the outset of the COVID-19 pandemic, Flint and neighboring Detroit represented one of the hardest hit regions outside of New York City; ongoing concerns of water quality have only exacerbated public health concerns [8]. Our analyses therefore sought to identify heterogeneity in longitudinal patterns of COVID-19 case diagnoses across Genesee County during different points in the pandemic. We then sought to identify community-level predictors that indicated what drove varying levels of COVID-19 cases during two

Conflicts of interest: The authors declare no conflicts of interest, financial or otherwise.

\* Corresponding author: Richard Casey Sadler, PhD, MPH, Michigan State University, Flint, MI

E-mail address: [sadler@msu.edu](mailto:sadler@msu.edu) (R.C. Sadler).

major waves of the virus observed in Michigan during the first year of the pandemic.<sup>1</sup>

## Methods

### Data

Data for this study were provided by the Genesee County Health Department. Positive COVID-19 diagnoses—based on detection of SARS-CoV-2 RNA in a clinical or autopsy specimen using a molecular amplification test—were compiled by the health department and identified by the day in which the case was diagnosed. Our team then geocoded these cases to the residential address and aggregated at the census block group (CBG) level ( $n = 373$ ). This level of aggregation was chosen because it provided the largest  $N$  while still enabling masking of all individuals inhabiting the geographic area in the sample and allowing us to link data to census demographic data. All other data were taken from estimates from the American Community Survey [5] or computed from data provided by the Genesee County Health Department.

### Measures

Our main dependent variable examined in analyses was a count of the number of COVID-19 diagnoses identified within each CBG, aggregated into 2 week testing intervals. Each CBG in the sample then had a rolling count of the number of COVID-19 diagnoses during each of these 2 week testing intervals beginning March 1, 2020 and concluding January 17, 2021.

Our three primary predictor variables included gender composition, racial composition, and percent of seniors. Gender composition was measured by the proportion of male to female residents living in each CBG. This was calculated by dividing the number of male residents by the total number of residents in a CBG. Because Genesee County is majority White (65.49%) but heavily segregated [9], a race variable corresponding to the ratio of White residents compared to all other residents in each CBG was computed by dividing the number of White residents by the number of total residents in each CBG. The proportion of residents living in each CBG 65 years of age or older was computed by dividing the number of senior residents by the total number of residents living in each CBG.

Additional variables controlled for in analyses included economic distress, distance to nearest testing site, population density, and total residents. Economic distress is measured by an unweighted sum of z-scores of percent impoverished, unemployed, less than high school education, and single parent households for each CBG (as used in past work in Flint, see [10]). Because access to testing may influence the likelihood that residents living within a given CBG actually get tested and receive a COVID-19 diagnosis, distance to the nearest testing site for each CBG was controlled for in analyses. This distance was measured in miles and mapped using network analysis in ArcGIS.<sup>2</sup> Population density was calculated as the number of residents per square mile for each CBG.

<sup>1</sup> The identification of the existence of two major waves of COVID-19 cases during the first year of the pandemic is based on several empirical points. First, there was an initial increase in cases during the very early days of the pandemic followed by a decline and plateau throughout the summer of 2020 (Wave 1). A surge in cases was then observed in late 2020 that can be observed by examining the raw case counts (Wave 2). While the magnitude of the case surges comprising these waves varied, their existence can be confirmed by visual examination of the raw daily case count data. The existence of the two waves is even more apparent when examining the raw daily death count data (State of Michigan, 2021).

<sup>2</sup> Two separate variables were used to measure distance, one for each COVID-19 wave. This was done because the number of testing sites in Genesee County increased between the first and second waves, changing the distance to the nearest testing site for many CBGs.

Total number of residents for each CBG was included in addition to population density because of the relatively large standard deviation in total residents for CBGs (Standard deviation = 588.725) for the range (Range = 0–3635).

### Analytic strategy

The analytic strategy utilized in this study proceeded in several phases. In a preliminary assessment, we use relative risks (RRs) to identify disparities in COVID-19 cases over time in Flint and Genesee County. The first phase of analyses entailed the use of group-based trajectory modeling (GBTM) to identify latent patterns of change and continuity in COVID-19 diagnoses identified at the CBG level. This entails the process of iteratively fixing of different numbers of polynomial functions of varying complexity to a set of longitudinal response data. Units of analysis are then assigned membership to a given trajectory group in the model that best approximates their own individual response trajectory. Nested model fit is then determined using Bayesian Information Criteria (BIC) to identify the best fitting model. Nagin also outlines several other criteria that a best fitting model should also meet if it is to be selected [11]. Two models were estimated: first wave and second wave.

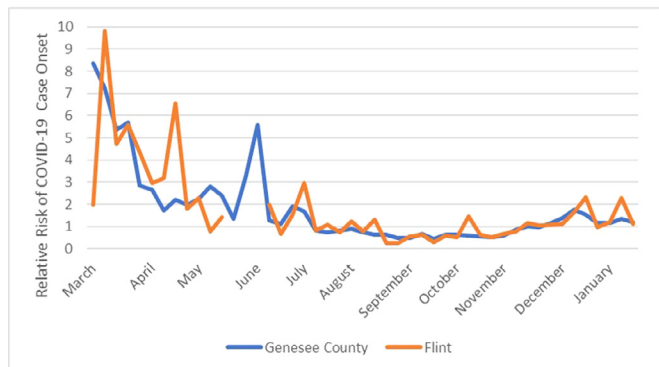
We split all analyses into two sections based on assessing the first large wave of COVID-19 cases and the second large wave of cases. This choice was made for several reasons. First, the data available to the research team spanned 25 total data points. As modeling evolved across time, nuance that was previously observed was lost, as case counts in the latter portion of the study period dwarfed earlier case counts; the model addressed this by focusing on larger frequencies observed later. The problem here is that the first surge in COVID-19 cases in the very early days of the pandemic are not properly described visually in the full model, despite the fact that the first wave was an empirical reality confirmed in the raw daily case counts in the full Michigan data and the Genesee County data [12]. Second, and relatedly, by the second wave, testing scope and accuracy had improved considerably. This means that although case counts are considerably higher in the second wave compared to the first, this is likely due in part to better testing. As such, it makes sense to examine waves separately and omit analysis of the relative lull in cases that was observed during summer 2020 in Michigan [12]. Similar decisions based on testing differences have been observed in prior work, like that focused on the HIV pandemic [13].

The first GBTM analyses examined the range of COVID-19 case data beginning on March 1, 2020 and ending on June 20, 2020. These cut-points were chosen because March 1, 2020 was the first available data, while the testing period ending June 20, 2020 is the lowest case count point of the first wave, with 90% of CBGs reporting zero cases. The second GBTM analyses focused on the dates ranging from August 16, 2020 until January 17, 2021. The starting point of August 16, 2020 was chosen because this was a low point at the beginning of the second wave where around 70% of CBGs reported zero cases; this percentage decreased considerably to around 40% at the next testing interval and continued to steadily decline through the crest of the second wave. The end point of January 17, 2021 was chosen because data saturation was reached at this point, as the addition of testing interval data points did not meaningfully alter findings.

The next phase of analyses entailed the use of regression modeling to identify predictors of trajectory group assignment. Logistic and multinomial regression were ultimately utilized in analyses based on the number of trajectory groups that were identified. Logistic regression is used when there is a binary dependent variable and coefficients are described as odds ratios (OR). The ORs indicate the predicted change in the odds of belonging in the 1 group

**Table 1**  
Descriptive Statistics

	Mean/Proportion	Standard Deviation	Minimum	Maximum
COVID-19 Cases Pooled Across Waves	2.016	4.137	0	65
Proportion of Male Residents	.483	.071	.210	.836
Proportion of White Residents	.655	.340	0	1
Distress Index	1.327	3.725	-4.500	18.633
Proportion of Residents Aged 65 Years or Older	.164	.084	0	.502
Population Density	2.267	2.069	0	9.441
Total Number of Residents	1097.483	587.969	0	3635
Distance to Nearest Testing Site (First Wave)	5.598	4.523	.087	18.724
Distance to Nearest Testing Site (Second Wave)	4.843	4.428	.173	18.724



**Figure 1.** Relative Risk of COVID-19 Cases, African-Americans vs. White/Other Races in Flint and Genesee County, MI.

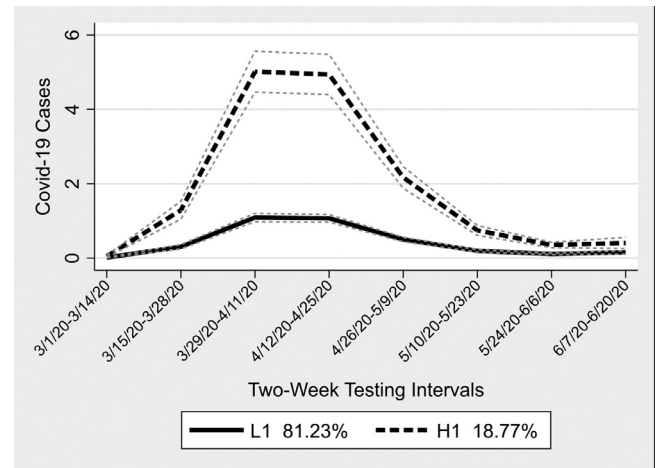
of a 0–1 coded dependent variable based on a one-unit change in an independent variable of interest. Multinomial logistic regression entails the omission of one trajectory group in the model, which then is examined in comparison to all other groups in the model in terms of how variation in independent variables of interest predict risk of assignment to the other groups relative to this excluded group. Coefficients are described in the form of relative risk ratios (RRR), indicating the predicted difference in risk of assignment to given group, relative to assignment to the omitted group, based on a one-unit change in an independent variable of interest.

**Results**

To provide us with an initial impression of race- and place-related disparities, we began our analyses by computing RRs of contracting COVID-19 by week for African-Americans versus all other races. **Figure 1** highlights a disparity, such that African-American residents were on average 3.3 times more likely to be diagnosed with COVID-19 than all other races from the beginning of the time period through the end of June 2020. From July to November, they actually experience a lower RR, at 0.75 in Flint and 0.88 in Genesee County. In December 2020 and January 2021, meanwhile, a disparity reappears, as they have RRs of 1.36 and 1.52 for Flint and Genesee County.

**Table 1** provides descriptive statistics for variables included in analyses. **Tables 2** and **3** provide regression results for the first and second wave models.

The first GBTM analyses examined only the first wave of the COVID-19 pandemic. A two-group model was selected as providing best fit. While a three-group model provided better fit based on BIC, the third group identified had a very small number of CBGs assigned to it (n = 6) and this precluded any useful analysis. Beyond that, this group greatly resembled the other two groups in the model, but had higher case counts. Selection of a two-group model resulted in combining the two higher case count groups



**Figure 2.** First Wave of COVID-19 Cases Aggregated by Census Block Group Trajectory Model.

into a single trajectory group, but group shape remained essentially the same. **Figure 2** depicts these groups, identified as the Low (Wave 1) and High (Wave 1); or L1 and H1, respectively. The groups demonstrate increases and peaks in COVID-19 case counts during April and May 2020 at varying levels with rank order in their case count peaks corresponding to their names. The CBGs in the sample were assigned membership to each trajectory group, with 81.23% assigned membership to the L1 group and 18.77% assigned to the H1 group. Both groups in the model were characterized by a cubic polynomial function.

Logistic regression analyses were conducted on the first trajectory model. The H1 group was the one category for the dependent variable and the L1 group was the 0 category. Results indicated that having a lower proportion of White residents increased odds of belonging to the H1 group (OR = 0.024). Having more total residents also was associated with increased odds of belonging to the H1 group (1.002).

**Figure 3** describes a map plotted and coded based on first wave trajectory group assignment. Many of the CBGs appearing in the high group appear in predominately African-American neighborhoods in the City of Flint and Flint Township, but are also dispersed throughout the county. This somewhat belies the reality of the individual data within these groups, which more strongly suggests a racial disparity in group assignment.

The second GBTM analyses examined the second wave of the pandemic. **Figure 4** provides visual depiction of these groups. Results indicated that a three-group model once again provided best fit to the data, with similar patterns of Low (Wave 2), Moderate (Wave 2), and High (Wave 2) groups identified; or L2, M2, and H2 groups, respectively. These groups all demonstrated rises, peaks, and falls in case counts across the second wave in rank order corresponding to their names. All CBGs in the sample were assigned

**Table 2**  
Logistic Regression Model of Covariate Effects on Odds of Assignment to the H1 Trajectory Group Relative to Assignment to the L1 Trajectory Group in Odds Ratios (OR) (n = 373)

	OR	P-Value	95% Confidence Interval	
Proportion of Male Residents	.711	.883	.008	65.850
Proportion of White Residents	.024	<i>P</i> < .001	.006	.107
Distress Index	.896	.091	.789	1.018
Proportion of Residents Aged 65 Years or Older	25.706	.113	.462	1429.721
Population Density	1.015	.869	.849	1.214
Total Number of Residents	1.002	<i>P</i> < .001	1.002	1.003
Distance to Nearest Testing Site	.999	.980	.908	1.099
Constant	.117	.123	.008	1.786

**Table 3**  
Multinomial Logistic Regression Model of Covariate Effects on Relative Risk of Assignment to the H2 Trajectory Group in Relative Risk Ratios (RRR) (n = 373)

	RRR	p-Value	95% Confidence Interval	
<i>H2 (omitted)</i>	—	—	—	
<i>L2</i>				
Proportion of Male Residents	.124	.704	≥0.001	5809.335
Proportion of White Residents	2.830	.506	.132	60.550
Distress Index	1.658	.001	1.234	2.229
Proportion of Residents Aged 65 Years or Older	.401	.812	≥0.001	732.440
Population Density	1.547	.037	1.026	2.331
Total Number of Residents	.991	<i>P</i> < .001	.989	.993
Distance to Nearest Testing Site	1.003	.969	.872	1.153
Constant	46,151.230	.001	79.036	2690.000
<i>M2</i>				
Proportion of Male Residents	.124	.704	≥0.001	2420.277
Proportion of White Residents	5.464	.234	.334	89.433
Distress Index	1.342	.030	1.029	1.751
Proportion of Residents Aged 65 Years or Older	3.735	.684	.007	2146.220
Population Density	1.353	.114	.930	1.966
Total Number of Residents	.996	<i>P</i> < .001	.995	.997
Distance to Nearest Testing Site	.991	.866	.892	1.101
Constant	372.302	.041	1.271	109,020.600

to a given trajectory group, with 41.29% assigned to L2 group (n = 161), 43.16% assigned to the M2 group (n = 154), and 15.55% assigned to the H2 group (n = 58). All groups in the model were characterized by a cubic polynomial function.

Multinomial logistic regression was utilized to analyze predictors of trajectory group assignment for this second wave model. The H2 group was designated as the omitted reference group. Results indicated that greater economic distress predicted greater risk of assignment to both the L2 and M2 groups, relative to the H2 group (L2 RRR=1.342; M2 RRR=1.658). Having fewer total residents also predicted greater risk of assignment to both the L2 and M2 groups, relative to the H2 group (L2 RRR=0.996; M2 RRR=0.991). Greater population density also predicted greater risk of assignment to the M2 group relative to the H2 group (RRR=1.547).<sup>3</sup>

Figure 5 describes a map plotted and coded based on second wave trajectory group assignment. In contrast to the first wave map in Figure 3, a very strong spatial pattern is present such that a low trajectory group is clearly visible in the City of Flint. Almost all of the CBGs appearing in the high group, meanwhile, appear in Genesee County’s suburbs, which are predominately White and have a lower poverty rate. This provides further evidence that an initial racial disparity in COVID-19 cases was eliminated by the second wave.

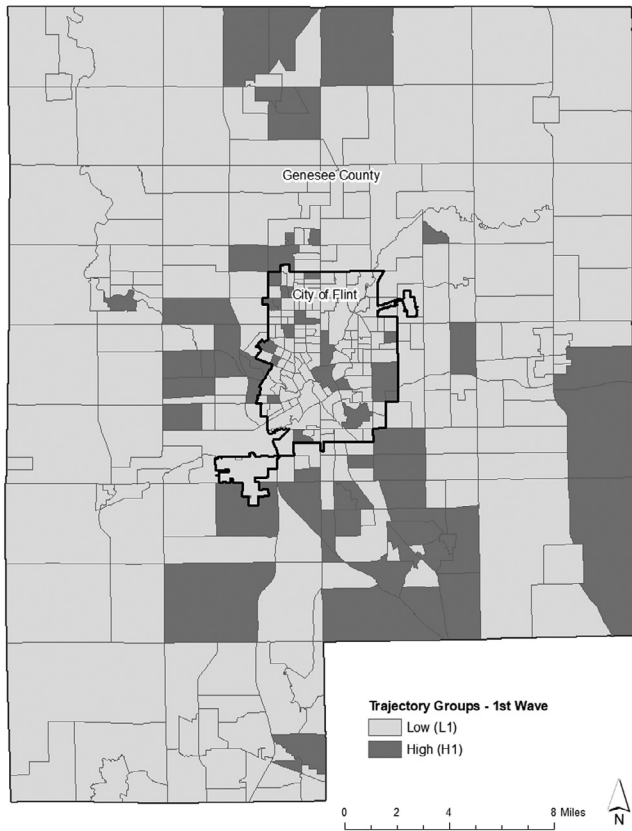
<sup>3</sup> Sensitivity analyses examined the ratio of African-American residents in each CBG as a predictor of trajectory group assignment, rather than White residents. Results generally were similar, with having a greater ratio of African-American residents predicting higher odds of assignment to the H1 group in the first wave and race remaining a nonsignificant predictor for the second wave.

## Discussion

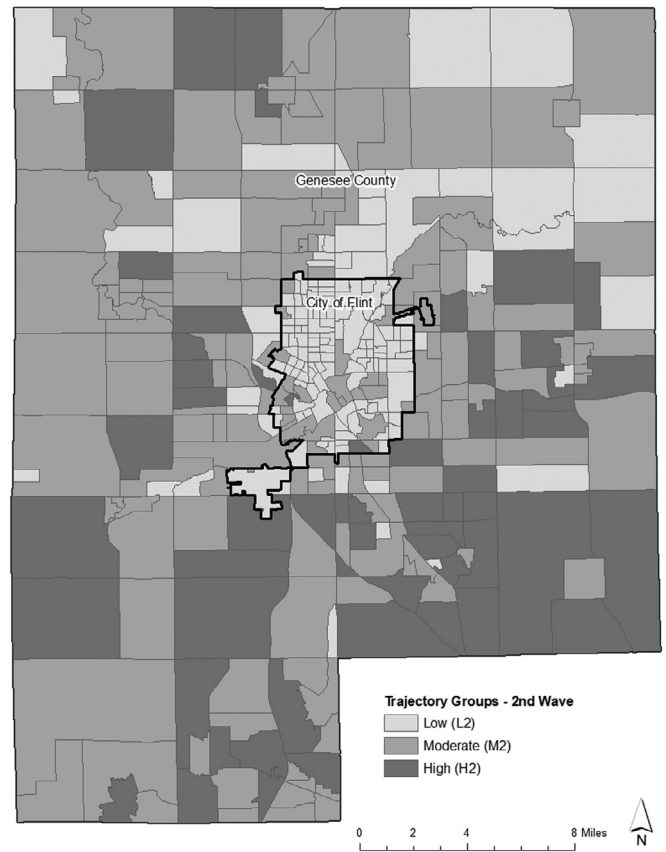
Our first key finding—that a racial disparity existed in COVID-19 case rates during the first wave in spring 2020—precipitated the subsequent work identifying latent attributes driving COVID-19 infection. Analyses indicated that the two waves of COVID-19 cases observed in Genesee County during the first year of the pandemic were driven by surges in communities with different demographic compositions. In the first wave (supporting our preliminary results), the highest case counts were seen in the City of Flint and surrounding areas, which are marked by higher socioeconomic distress and thus greater vulnerability to severe COVID-19 cases [14–15]. Economic distress was also a significant factor at the *P* ≤ .10 level. While not significant at the traditional threshold, sensitivity analyses indicated that removal of the total number of residents variable resulted in this becoming a significant predictor, indicating that more distressed CBGs were more likely to be in the H1 group. This indicates that economic distress must be considered when understanding drivers of the first wave. The highest case counts of the second wave, meanwhile, were observed among the more affluent and more populous CBGs in Genesee County.

Important implications can be drawn from the findings of our study. Foremost, it offers insights into the way the COVID-19 pandemic became politically charged, with suburban, rural, and Whiter areas being more likely to eschew masking and physical distancing mandates [16–18]. While predominately African-American Flint was disproportionately affected in the first wave, the second wave was more visible in southern Genesee County.

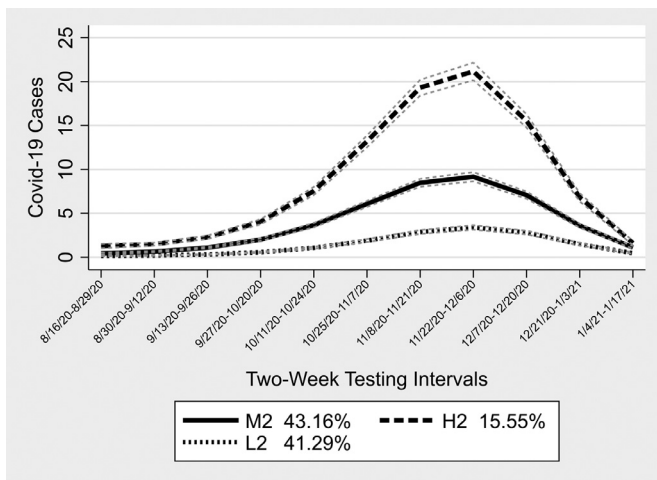
Second, our partnership with the Genesee County Health Department allowed this information to be continuously provided to



**Figure 3.** Trajectory Groups for the First Wave of COVID-19 Cases, Genesee County, March 1, 2020 to June 20, 2020.



**Figure 5.** Trajectory Groups for the Second Wave of COVID-19 Cases, Genesee County, August 16, 2020 to January 17, 2021.



**Figure 4.** Second Wave of COVID-19 Cases Aggregated by Census Block Group Trajectory Model.

stakeholders and the general public via the creation of a COVID Disparities Dashboard [19]. Through this work and related advocacy, we highlighted the elimination of a racial disparity in COVID caseloads prior to the onset of the second wave [20]; the lack of a racial disparity in the second wave supports this assertion.

One limitation is an inability to parse out reasons for the differences in predictors of the H1 and H2 groups between waves. While race and affluence appear to be predictors, underlying reasons why these differences existed is beyond the scope of this study. More affluent CBGs driving high case counts in the second wave model provides an example here. As noted, lax social dis-

tancing and other behavioral differences among residents of more affluent CBGs may have led to high case counts there. Alternatively, residents of more economically distressed neighborhoods may have been getting tested at lower rates because of precarious employment and/or finances. A positive test may mean a quarantine that would lead to further economic peril that these individuals could not afford, thus their case counts would be artificially deflated. These competing mechanisms provide an example of these data limitations and highlight the need for future research to interrogate these points further so that public health interventions, policy, and outreach can be more effectively targeted.

**IRB approval**

The Genesee County Health Department and Michigan State University Institutional Review Board approved this study.

**References**

- [1] Brammer S, Branicki L, Linnenluecke MK. COVID-19, societalization, and the future of business in society. *Academy of Management Perspectives* 2020;34(4):493–7.
- [2] Chakraborty I, Maity P. COVID-19 outbreak: migration, effects on society, global environment and prevention. *Science of the Total Environment* 2020;728:138882.
- [3] Franch-Pardo I, Napoletano BM, Rosete-Verges F, Billa L. Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment* 2020;739:140033.
- [4] Zahran, S, McElmurry, S.P, Kilgore, P.E., Mushinski, D., Press, J., Love, N.G., et al. (2018). Assessment of the Legionnaires' disease outbreak in Flint, Michigan. *Proceedings of the National Academy of Sciences*, 115(8), E1730–9.
- [5] Manson S, Schroeder J, Van Riper D, Ruggles S. IMPUS national historical geographic information system: version 12.0 [database]. Minneapolis: University of Minnesota; 2017. p. 39.

- [6] Hanna-Attisha M, LaChance J, Sadler RC, Champney Schnepf A. Elevated blood lead levels in children associated with the Flint drinking water crisis: a spatial analysis of risk and public health response. *Am J Public Health* 2016;106(2):283–90.
- [7] Sadler RC, Highsmith AR. Rethinking Tiebout: the contribution of political fragmentation and racial/economic segregation to the Flint water crisis. *Environmental Justice* 2016;9(5):143–51.
- [8] Hyde, K. (2020). Residential water quality and the spread of COVID-19 in the United States. Available at SSRN 3572341.
- [9] Sadler RC, Lafreniere DJ. Racist housing practices as a precursor to uneven neighborhood change in a post-industrial city. *Hous Stud* 2017;32(2):186–8.
- [10] Sadler RC, LaChance J, Hanna-Attisha M. Social and built environmental correlates of predicted blood lead levels in the Flint water crisis. *Am J Public Health* 2017;107(5):763–9.
- [11] Nagin D. *Group-based modeling of development*. Cambridge, Mass: Harvard University Press; 2005.
- [12] State of Michigan (2021). Michigan.gov Data Dashboard. Available at: [https://www.michigan.gov/coronavirus/0,9753,7-406-98163\\_98173--,00.html](https://www.michigan.gov/coronavirus/0,9753,7-406-98163_98173--,00.html).
- [13] Johnson CC, Fonner V, Sands A, Ford N, Obermeyer CM, Tsui S, et al. To err is human, to correct is public health: a systematic review examining poor quality testing and misdiagnosis of HIV status. *J Int AIDS Soc* 2017;20:21755.
- [14] Sadler RC, Furr-Holden D. The epidemiology of opioid overdose in Flint and Genesee County, Michigan: implications for public health practice and intervention. *Drug Alcohol Depend* 2019;204:107560.
- [15] Patel JA, Nielsen FBH, Badiani AA, Assi S, Unadkat VA, Patel B, et al. Poverty, inequality and COVID-19: the forgotten vulnerable. *Public Health* 2020;183:110.
- [16] Hearne BN, Niño MD. Understanding How Race, Ethnicity, and Gender Shape Mask-Wearing Adherence During the COVID-19 Pandemic: evidence from the COVID Impact Survey. *J Racial Ethn Health Disparities* 2021:1–8 <https://link.springer.com/content/pdf/10.1007/s40615-020-00941-1.pdf>.
- [17] Rader B, White LF, Burns MR, Chen J, Brilliant J, Cohen J, et al. Mask-wearing and control of SARS-CoV-2 transmission in the USA: a cross-sectional study. *The Lancet Digital Health* 2021;3(3):e148–57.
- [18] Makridis, C., & Rothwell, J.T. (2020). The real cost of political polarization: evidence from the COVID-19 pandemic. Available at SSRN 3638373.
- [19] Genesee County Health Department (2021). Genesee County COVID Disparities Dashboard. Available at: <https://msugis.maps.arcgis.com/apps/opsdashboard/index.html#/c0373cd7238b4103aecbbdeec9f4e62f>.
- [20] ABC12 News (2020). Flint eliminates COVID-19 disparity among African-Americans. Available at: <https://www.abc12.com/2020/11/14/flint-eliminates-covid-19-disparity-among-african-americans>.