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.
The association of voter turnout with county-level coronavirus disease 2019 occurrence early in the pandemic

W. Dana Flanders, MD, DSc a, *, William D. Flanders, PhD, MS b, Michael Goodman, MD, MPH a

a Department of Epidemiology, Rollins School of Public Health, Emory University, Atlanta, GA
b Wisconsin Institute for Law & Liberty, Milwaukee, WI

ARTICLE INFO

Article history:
Received 8 May 2020
Accepted 23 June 2020
Available online 1 July 2020

Keywords:
COVID-19
Risk factors
Social distancing
Voting

ABSTRACT

Purpose: The ongoing coronavirus disease 2019 (COVID-19) severely impacted both health and the economy. Absent an effective vaccine, preventive measures used, some of which are being relaxed, have included school closures, restriction of movement, and banning of large gatherings. Our goal was to estimate the association of voter turnout with county-level COVID-19 risks.

Methods: We used publicly available data on voter turnout in the March 10 primary in three states, COVID-19 confirmed cases by day and county, and county-level census data. We used zero-inflated negative binomial regression to estimate the association of voter turnout with COVID-19 incidence, adjusted for county-level population density and proportions: over age 65 years, female, Black, with college education, with high school education, poor, obese, and smokers.

Results: COVID-19 risk was associated with voter turnout, most strongly in Michigan during the week starting 3 days postelection (risk ratio, 1.24; 95% confidence interval, 1.16–1.33). For longer periods, the association was progressively weaker (risk ratio 0.98–1.03).

Conclusions: Despite increased absentee-ballot voting in the primary, our results suggest an association of voter turnout in at least one state with a detectable increase in risks associated with and perhaps due to greater exposures related to the primary.

© 2020 Elsevier Inc. All rights reserved.

Introduction

Although coronavirus-induced epidemics occur periodically [1], the coronavirus disease 2019 (COVID-19) has had especially severe public health impact [2,3]. Compared with earlier coronavirus epidemics, such as those caused by severe acute respiratory syndrome (SARS) and the Middle East respiratory syndrome, a greater proportion of COVID-19 cases are infectious while still asymptomatic [4,5]. As the result, unlike SARS and Middle East respiratory syndrome, which were mainly associated with nosocomial spread, the Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), which causes COVID-19, is more easily transmitted in the community [6,7].

Based on experience with other infectious disease epidemics, community transmission typically displays pronounced spatial heterogeneity that depends on two key factors: where people live and how they move or gather [8,9]. The latter consideration is the main justification for social distancing measures, such as school and university closures, cancellation of planned events, and restriction of movement [10–13]. Despite these measures, population gatherings, including family events such as funerals and birthdays, facilitate SARS-CoV-2 transmission in various settings [14,15].

Although the current wave of the COVID-19 pandemic is expected to subside, lower level transmission will likely continue, and a second wave in the fall is possible [16]. If so, it could directly affect the presidential elections, scheduled for November 3, 2020 [17].

In planning for the November election, it may be helpful to consider the recent experience in the states that held primary elections on March 10, 2020, as planned. Data from these states offer an opportunity to investigate the potential impact of the reduction in social distancing that might be caused by in-person voting. The impact, if any, may inform the administrative and logistical measures that need to be considered for November.

To estimate the impact of elections on county-level COVID-19 incidence, we considered voter turnout differences in the March 10
primary, by county, across the three states for which the required information was available. The purpose of the current analysis is to investigate whether the increase in cases that was seen in all states after the election was greater in counties with higher voter turnout after accounting for other relevant county-level population characteristics.

Methods

Cases

We obtained information on confirmed COVID-19 cases by county from two sources—USA Facts [18] (Table 1) and the Johns Hopkins University Center for Science and Engineering project [19] (JHU-CSSE; Table 1). Both sources report confirmed case counts by county and day, using data from the U.S. Centers for Disease Control and Prevention (CDC) and from state and local governments. (More recently, CDC began reporting both confirmed and probable cases, but these changes do not affect the present report.) Although numbers from USAFacts and JHU-CSSE differ (see Tables 3–6), the results using either source led to the same conclusions, and therefore, the main analyses use the JHU-CSSE data.

We restricted analyses to the three states with primary elections on March 10, 2020, and with the required information as of May 05, 2020: Michigan, Mississippi, and Missouri. Although they also held a primary on March 10, we excluded Idaho because not all counties had completed tallies; Washington because voters cast ballots by mail; and North Dakota because it has caucuses, not primary elections.

Voter turnout and covariates

Information on voter turnout, used as a surrogate of in-person voting in Missouri and Mississippi, was gathered from CNN’s Voter turnout and covariates elections. We restricted analyses to the three states with primary elections on March 10, 2020, and with the required information as of May 05, 2020: Michigan, Mississippi, and Missouri. Although they also held a primary on March 10, we excluded Idaho because not all counties had completed tallies; Washington because voters cast ballots by mail; and North Dakota because it has caucuses, not primary elections.

Table 1: Sources of data (publicly available)

| Variable | Source | Data |
|----------|--------|------|
| Outcome: COVID Case/Death Count | John Hopkins University, the Center for Systems Science and Engineering [https://coronavirus.jhu.edu/map.html](https://coronavirus.jhu.edu/map.html) | Daily death and case counts at the county level |
| Outcome: COVID Case/Death Count | USA Facts Corona Virus Live Map [https://usafacts.org/visualizations/coronavirus-covid-19-spread-map](https://usafacts.org/visualizations/coronavirus-covid-19-spread-map) | Daily death and case counts at the county level |
| Independent Variable: Turnout Michigan | CNN Michigan Primary 2020 [https://www.cnn.com/election/2020/state/michigan?xid=ec_state_michigan_d](https://www.cnn.com/election/2020/state/michigan?xid=ec_state_michigan_d) | Primary election results for the Democratic and Republican Presidential primaries at the county level |
| Independent Variable: Turnout Mississippi | CNN Mississippi Primary 2020 [https://www.cnn.com/election/2020/state/mississippi](https://www.cnn.com/election/2020/state/mississippi) | Primary election results for the Democratic and Republican Presidential primaries at the county level |
| Independent Variable: Turnout Washington | CNN Washington Primary 2020 [https://www.cnn.com/election/2020/state/washington](https://www.cnn.com/election/2020/state/washington) | Primary election results for the Democratic and Republican Presidential primaries at the county level |
| Independent Variable: Turnout Missouri | CNN Missouri Primary 2020 [https://www.cnn.com/election/2020/state/missouri](https://www.cnn.com/election/2020/state/missouri) | Primary election results for the Democratic and Republican Presidential primaries at the county level |
| Controls: Demographic Data | United States Census Bureau Quick Facts [https://www.census.gov/quickfacts/fact/table/US/PST045219](https://www.census.gov/quickfacts/fact/table/US/PST045219) | Poverty, race, density, female variables (2019 census estimates) |
| Controls: Health Data | CDC Behavioral Risk Surveillance System Surveys [https://www.cdc.gov/brfss/index.html](https://www.cdc.gov/brfss/index.html) | Smoking and obesity data at the county level (BRFSS 2018) |

Statistical analyses

To estimate the association between voter turnout on March 10 and county-level COVID-19 infection risk, we considered the period during which excess cases, if any, would be expected to occur. Therefore, we considered published estimates concerning the incubation period, time from infection to symptom onset and the interval from onset of symptoms to the development of dyspnea or hospitalization. These estimates are relevant because CDC, the main information source for JHU-CSSE (and USAFacts), linked some cases with date reported, not necessarily with symptom onset date. The estimates of the median incubation period range from about 5 [20–24] to 6.5 or more days [25,26]. Lauer et al. [22] estimated that only 2.5% of people would develop symptoms sooner than 2.6 days after infection, and 97.5% of people would develop symptoms within 11.5 days of infection. In a study by Wu et al. [25], the 97.5 percentile of the COVID-19 incubation-period distribution was about 12.5 days. Zhou et al. estimated a median time of 7 days (interquartile range [IQR] 4–9) from symptom onset to dyspnea and 11 (IQR 8–14) days [27] to hospital admission, and Chen reported a median time from symptom onset until the presentation of 4 days (IQR 4–7).

Based on this information, in the main analyses, we considered COVID-19 cases reported through March 10 to be “prevote,” and considered cases reported on March 13 and after as “postvote.” We calculated case counts for each of a series of risk periods, ranging in length from 5 to 12 days, as indicated in Tables 3–10. We chose these risk periods to allow for the development of symptoms (incubation period) plus time for obtaining and reporting a test result after symptom onset. Collectively, the risk periods considered (Tables 3–10) could include poll-associated cases that were tested as soon as 3–4 days and as long as 15 days after voting. These periods allow for the median incubation period (5–6 days) plus up to an additional 10 days for testing and reporting, because, for part of this time, CDC apparently reported counts by reported date (not necessarily the symptom onset date) [28]. Cases reported on March 11 and March 12 were not included in the analyses because they could have become infected prevote. In sensitivity analyses, we considered alternative risk periods to define “postvote” cases as described under “alternative” outcomes.

To control for potential confounding of the voter turnout-COVID-19 association, we generated a list of variables that were suspected risk factors for COVID-19 infection. The list included state and county-level population characteristics: density, prevote COVID-19 risk, percent female, percent Black, percent older than 65 years, percent “poor” (living below the federal poverty line),...
percent obese, percent smokers, percent with a college education, and percent with a high school education. From this list, we selected a priori, the most important risk factors for COVID-19 that may also influence county-level voter turnout; these were, in addition to state, county population density, percent female, percent Black, percent older than 65 years, and percent poor.

Descriptive statistics were computed to characterize the distribution of the counts of postpolling cases, voter turnout, and covariates. Simple linear regression was used to describe the association of voter turnout with the postvote COVID-19 risk (Fig. 1). The primary analyses used a zero-inflated negative binomial regression model to account for a more-than-expected number of counties with a 0 count and because of improved fit compared with Poisson models (based on corrected Akaike Information Criterion and Pearson’s χ²). The logarithm of the population size was an offset in all models.

### Alternative outcomes

We reasoned that if the association of voter turnout with COVID-19 occurrence in the postvote period was attributable to uncontrolled confounding, then that association should persist even if we redefined the outcome as cases occurring in risk periods (e.g., in early April) that did not overlap substantially with the time interval of interest (the incubation period plus some allowance for testing and reporting). That is, if confounding (e.g., behavioral patterns, pre-existing disease, and other risk factors) explained the observed association, we would expect the association to persist, likely not substantially weaker, even long after the polling. In contrast, if the

---

**Table 2**

Summary statistics of key variables

**Michigan (n = 83)**

| Variable          | Postvote cases | Postvote cases | Cases per 100,000 | Cases per 100,000 | Turnout | Age | Black | Female | Density | Poverty | Prevalence of smoking | Prevalence of obesity |
|-------------------|----------------|----------------|--------------------|--------------------|---------|-----|-------|--------|---------|---------|-----------------------|-----------------------|
| Mean              | 15.93          | 4.30           | 2.43               | 1.64               | 0.15     | 0.22| 0.04  | 0.50   | 189.16  | 0.14    | 0.18                  | 0.34                  |
| SD                | 80.13          | 20.12          | 5.64               | 3.17               | 0.04     | 0.05| 0.06  | 0.02   | 409.55  | 0.04    | 0.02                  | 0.04                  |
| Median            | 0.00           | 0.00           | 0.00               | 0.00               | 0.14     | 0.20| 0.01  | 0.50   | 59.60   | 0.14    | 0.18                  | 0.34                  |

**Missouri (n = 115)**

| Variable          | Postvote cases | Postvote cases | Cases per 100,000 | Cases per 100,000 | Turnout | Age | Black | Female | Density | Poverty | Prevalence of smoking | Prevalence of obesity |
|-------------------|----------------|----------------|--------------------|--------------------|---------|-----|-------|--------|---------|---------|-----------------------|-----------------------|
| Mean              | 0.85           | 0.94           | 0.83               | 0.82               | 0.14     | 0.20| 0.04  | 0.50   | 131.30  | 0.16    | 0.21                  | 0.33                  |
| SD                | 2.87           | 3.60           | 1.91               | 1.92               | .03      | 0.04| 0.06  | 0.02   | 523.10  | 0.05    | 0.02                  | 0.04                  |
| Median            | 0.00           | 0.00           | 0.00               | 0.00               | 0.13     | 0.19| 0.01  | 0.50   | 31.00   | 0.16    | 0.20                  | 0.33                  |

**Mississippi (n = 82)**

| Variable          | Postvote cases | Postvote cases | Cases per 100,000 | Cases per 100,000 | Turnout | Age | Black | Female | Density | Poverty | Prevalence of smoking | Prevalence of obesity |
|-------------------|----------------|----------------|--------------------|--------------------|---------|-----|-------|--------|---------|---------|-----------------------|-----------------------|
| Mean              | 3.02           | 3.02           | 7.93               | 7.93               | 0.20     | 0.17| 0.42  | 0.51   | 62.24   | 0.23    | 0.2                   | 0.38                  |
| SD                | 4.82           | 4.82           | 5.39               | 5.72               | 0.05     | 0.03| 0.21  | 0.02   | 64.03   | 0.07    | 0.03                  | 0.05                  |
| Median            | 1.00           | 1.00           | 9.72               | 5.34               | 0.19     | 0.17| 0.38  | 0.52   | 45.05   | 0.22    | 0.2                   | 0.38                  |

---

**Table 3**

RRs for all states combined, measuring the association of voter turnout with COVID-19 risk (we defined the outcome [case-count] using different at-risk periods and adjusted for population density and state)

| At-risk period used to define the case count | Analyses using COVID-19 case counts from USAFacts | Analyses using COVID-19 case counts from JHU-CSSE |
|--------------------------------------------|---------------------------------------------------|--------------------------------------------------|
| March 13 to March 17                        | n = 48; RR, 1.06; CI, 0.97–1.16; P = .20           | n = 90; RR, 1.10; CI, 1.02–1.19; P = .01         |
| March 13 to March 18                        | n = 71; RR, 1.05; CI, 0.97–1.14; P = .20           | n = 164; RR, 1.05; CI, 0.99–1.14; P = .22        |
| March 13 to March 19                        | n = 158; RR, 1.06; CI, 0.98–1.15; P = .13          | n = 319; RR, 1.07; CI, 1.01–1.15; P = .09        |
| March 13 to March 20                        | n = 271; RR, 1.07; CI, 1.01–1.15; P = .03          | n = 523; RR, 1.08; CI, 1.01–1.16; P = .02        |
| March 13 to March 21                        | n = 402; RR, 1.05; CI, 1.00–1.10; P = .05          | n = 687; RR, 1.07; CI, 1.01–1.14; P = .01        |
| March 13 to March 22                        | n = 554; RR, 1.05; CI, 1.00–1.10; P = .05          | n = 1304; RR, 1.06; CI, 1.01–1.12; P = .02       |
| March 13 to March 23                        | n = 704; RR, 1.04; CI, 0.99–1.08; P = .11          | n = 1668; RR, 1.05; CI, 1.00–1.09; P = .05       |
| March 13 to March 24                        | n = 932; RR, 1.01; CI, 0.97–1.05; P = .59          | n = 2234; RR, 1.02; CI, 0.98–1.07; P = .37       |
| March 13 to March 25                        | n = 1202; RR, 1.01; CI, 0.98–1.05; P = .46         | n = 2387; RR, 1.02; CI, 0.98–1.06; P = .24       |
| April 3 to April 7                          | n = 3804; RR, 1.02; CI, 0.99–1.06; P = .12         | n = 9305; RR, 1.01; CI, 0.98–1.05; P = .39       |
| April 3 to April 8                          | n = 4645; RR, 1.02; CI, 0.99–1.06; P = .15         | n = 10885; RR, 1.01; CI, 0.98–1.05; P = .38      |
| April 3 to April 9                          | n = 5463; RR, 1.02; CI, 0.99–1.05; P = .18         | n = 12401; RR, 1.01; CI, 0.98–1.04; P = .36      |
| April 3 to April 10                         | n = 6312; RR, 1.01; CI, 0.99–1.05; P = .36         | n = 14036; RR, 1.01; CI, 0.98–1.04; P = .65      |
| April 3 to April 11                         | n = 7119; RR, 1.02; CI, 0.99–1.05; P = .32         | n = 12815; RR, 1.01; CI, 0.98–1.04; P = .70      |

* Counts of confirmed COVID-19 cases downloaded from USAFacts and from GitHub managed by Johns Hopkins University Center for Systems Science and Engineering [see main text].

† Adjusted for logarithm of population density and state; the zero-inflation part of the model included the intercept only.
The results of each analysis are expressed as a risk ratio (RR) representing the average difference in postelection COVID-19 risk per 1% difference in voter turnout. RRs are accompanied by 95% confidence intervals (CIs) and corresponding P values. Regression diagnostics include assessing the correlation between the independent variables, residual analyses, sensitivity analyses, and identification of potentially influential points. Analyses were done using SAS version 9.4 (SAS Institute, Cary, NC) and R version 4.0.

To assess the sensitivity of results to model specification and residual confounding, we conducted two main types of sensitivity analyses (see Online Supplement). First, we used a different model in which we defined ordinal outcomes by grouping county-specific COVID-19 risks and assumed a multinomial distribution with a cumulative logit link. Second, we allowed for within-state correlation of county rates and controlled for additional covariates using a random-effects, zero-inflated negative binomial model.

### Results

COVID-19 counts obtained from the two sources, JHU-CSSE and USAFacts, were highly correlated (e.g., $r > 0.99$) but not identical. For example, from March 13 to March 23, 1668 cases were reported...
Table 6

| At-risk period used to define the case count | Analyses using COVID-19 case counts from USAFacts | Analyses using COVID-19 case counts from JHU-CSSE |
|--------------------------------------------|--------------------------------------------------|--------------------------------------------------|
| March 13 to March 17                       | $n = 20$; RR, 1.10; CI, 0.90–1.33; $P = .35$       | $n = 20$; RR, 1.09; CI, 0.89–1.33; $P = .35$ |
| March 13 to March 18                       | $n = 33$; RR, 0.93; CI, 0.78–1.10; $P = .39$       | $n = 33$; RR, 0.99; CI, 0.88–1.10; $P = .79$   |
| March 13 to March 19                       | $n = 49$; RR, 0.95; CI, 0.84–1.08; $P = .45$       | $n = 49$; RR, 0.96; CI, 0.85–1.09; $P = .51$   |
| March 13 to March 20                       | $n = 79$; RR, 1.00; CI, 0.91–1.09; $P = .95$       | $n = 79$; RR, 1.00; CI, 0.91–1.09; $P = .97$   |
| March 13 to March 21                       | $n = 139$; RR, 0.98; CI, 0.92–1.04; $P = .53$      | $n = 97$; RR, 1.00; CI, 0.91–1.09; $P = .97$   |
| March 13 to March 22                       | $n = 206$; RR, 0.99; CI, 0.93–1.05; $P = .71$      | $n = 206$; RR, 0.99; CI, 0.93–1.05; $P = .64$  |
| March 13 to March 23                       | $n = 248$; RR, 0.98; CI, 0.92–1.04; $P = .49$      | $n = 248$; RR, 0.98; CI, 0.92–1.04; $P = .49$  |
| March 13 to March 24                       | $n = 313$; RR, 0.96; CI, 0.91–1.02; $P = .18$      | $n = 319$; RR, 0.96; CI, 0.91–1.02; $P = .17$  |
| March 13 to March 25                       | $n = 375$; RR, 0.97; CI, 0.92–1.03; $P = .33$      | $n = 376$; RR, 0.97; CI, 0.92–1.03; $P = .32$  |
| April 3 to April 7                        | $n = 739$; RR, 1.00; CI, 0.96–1.04; $P = .90$      | $n = 738$; RR, 1.00; CI, 0.96–1.04; $P = .86$  |
| April 3 to April 8                        | $n = 827$; RR, 1.00; CI, 0.96–1.04; $P = .97$      | $n = 826$; RR, 1.00; CI, 0.96–1.04; $P = .99$  |
| April 3 to April 9                        | $n = 1082$; RR, 1.00; CI, 0.97–1.04; $P = .85$     | $n = 1083$; RR, 1.00; CI, 0.97–1.04; $P = .82$ |
| April 3 to April 10                       | $n = 1289$; RR, 1.00; CI, 0.96–1.04; $P = .89$     | $n = 1292$; RR, 1.00; CI, 0.97–1.04; $P = .87$ |
| April 3 to April 11                       | $n = 1466$; RR, 1.00; CI, 0.97–1.04; $P = .82$     | $n = 1465$; RR, 1.00; CI, 0.96–1.04; $P = .85$ |

* Counts of confirmed COVID-19 cases downloaded from USAFacts and from GitHub managed by Johns Hopkins University Center for Systems Science and Engineering (see main text).
^ Adjusted for logarithm of population density; the zero-inflation part of the model included the intercept only.

Table 7

| At-risk period used to define the case count | Analyses using COVID-19 case counts from JHU-CSSE |
|--------------------------------------------|--------------------------------------------------|
| March 13 to March 18                       | $RR = 1.05$; CI, 0.97–1.15; $P = .25$           |
| March 13 to March 19                       | $RR = 1.06$; CI, 0.98–1.14; $P = .12$           |
| March 13 to March 20                       | $RR = 1.08$; CI, 1.01–1.15; $P = .03$           |
| March 13 to March 21                       | $RR = 1.07$; CI, 1.01–1.14; $P = .02$           |
| March 13 to March 22                       | $RR = 1.06$; CI, 1.01–1.12; $P = .03$           |
| March 13 to March 23                       | $RR = 1.04$; CI, 1.00–1.09; $P = .07$           |
| March 13 to March 24                       | $RR = 1.02$; CI, 0.98–1.07; $P = .31$           |
| March 13 to March 25                       | $RR = 1.02$; CI, 0.98–1.07; $P = .26$           |
| Alternative Outcomes                      |                                                   |
| April 3 to April 8                        | $RR = 1.00$; CI, 0.97–1.04; $P = .84$           |
| April 3 to April 9                        | $RR = 1.00$; CI, 0.97–1.04; $P = .82$           |
| April 3 to April 10                       | $RR = 1.00$; CI, 0.97–1.04; $P = .82$           |
| April 3 to April 11                       | $RR = 1.00$; CI, 0.97–1.03; $P = .84$           |
| April 3 to April 12                       | $RR = 1.00$; CI, 0.97–1.03; $P = .81$           |

* Counts of confirmed COVID-19 cases from GitHub managed by Johns Hopkins University Center for Systems Science and Engineering (see main text).
^ Adjusted for logarithm of population density, proportion female, proportion ages ≥65 y, proportion Black and proportion living below the poverty line, in count model. The zero-inflation model included the intercept only.
Table 9
RRs in Michigan, measuring the association of voter turnout with COVID-19 risk (we defined the outcome [case-count] using different at-risk periods and adjusted for population density, state and demographic covariates)

| At-risk period used to define the case count | Analyses using COVID-19 case counts from JHU-CSSE |
|--------------------------------------------|--------------------------------------------------|
| March 13 to March 18 RR = 1.14            | 95% CI, 0.96; 1.35; P = .13                      |
| March 13 to March 19 RR = 1.04            | 95% CI, 0.93; 1.16; P = .49                      |
| March 13 to March 20 RR = 1.02            | 95% CI, 0.94; 1.10; P = .71                      |
| March 13 to March 21 RR = 1.05            | 95% CI, 0.97; 1.12; P = .19                      |
| March 13 to March 22 RR = 1.07            | 95% CI, 1.00; 1.14; P = .03                      |
| March 13 to March 23 RR = 1.08            | 95% CI, 0.99; 1.18; P = .09                      |
| March 13 to March 24 RR = 1.07            | 95% CI, 0.99; 1.16; P = .09                      |
| March 13 to March 25 RR = 1.08            | 95% CI, 1.00; 1.16; P = .04                      |
| Alternative outcomes                      |                                                  |
| April 3 to April 8 RR = .97              | 95% CI, 0.91; 1.04; P = .36                      |
| April 3 to April 9 RR = .96              | 95% CI, 0.90; 1.03; P = .26                      |
| April 3 to April 10 RR = .96             | 95% CI, 0.90; 1.03; P = .27                      |
| April 3 to April 11 RR = .97             | 95% CI, 0.91; 1.04; P = .30                      |
| April 3 to April 12 RR = .97             | 95% CI, 0.91; 1.03; P = .38                      |

* Counts of confirmed COVID-19 cases downloaded from GitHub managed by Johns Hopkins University Center for Systems Science and Engineering (see main text).
1 Adjusted for logarithm of population density, proportion female, proportion aged ≥65 y, proportion Black and proportion living below the poverty line, in count model. The zero-inflation model included the intercept only.

Alternative outcomes

In the alternative outcome analyses, we defined postvoting cases as those occurring in several periods starting on April 3. Controlling for the same variables as in the a priori model, each of the RRs was close to 1.0 (bottom rows of Tables 7–9), and all were consistent with no association.

The results of the supplemental analyses were generally consistent with the main results (Online Supplement). For example, for the March 13 to March 20 risk period in Michigan using ordinal logistic regression, the odds ratio relating voter turnout and COVID-

Table 10
RRs in Mississippi, measuring the association of voter turnout with COVID-19 risk (we defined the outcome [case-count] using different at-risk periods and adjusted for population density, state and demographic covariates)

| At-risk period used to define the case count | Analyses using COVID-19 case counts from JHU-CSSE |
|--------------------------------------------|--------------------------------------------------|
| March 13 to March 18 RR = 1.14            | 95% CI, 0.96; 1.35; P = .13                      |
| March 13 to March 19 RR = 1.04            | 95% CI, 0.93; 1.16; P = .49                      |
| March 13 to March 20 RR = 1.02            | 95% CI, 0.94; 1.10; P = .71                      |
| March 13 to March 21 RR = 1.05            | 95% CI, 0.97; 1.12; P = .19                      |
| March 13 to March 22 RR = 1.07            | 95% CI, 1.00; 1.14; P = .03                      |
| March 13 to March 23 RR = 1.08            | 95% CI, 0.99; 1.18; P = .09                      |
| March 13 to March 24 RR = 1.07            | 95% CI, 0.99; 1.16; P = .09                      |
| March 13 to March 25 RR = 1.08            | 95% CI, 1.00; 1.16; P = .04                      |
| Alternative Outcomes                      |                                                  |
| April 3 to April 8 RR = .97              | 95% CI, 0.91; 1.04; P = .36                      |
| April 3 to April 9 RR = .96              | 95% CI, 0.90; 1.03; P = .26                      |
| April 3 to April 10 RR = .96             | 95% CI, 0.90; 1.03; P = .27                      |
| April 3 to April 11 RR = .97             | 95% CI, 0.91; 1.04; P = .30                      |
| April 3 to April 12 RR = .97             | 95% CI, 0.91; 1.03; P = .38                      |

* Counts of confirmed COVID-19 cases downloaded from GitHub managed by Johns Hopkins University Center for Systems Science and Engineering (see main text).
1 Adjusted for logarithm of population density, proportion female, proportion aged ≥65 y, proportion Black and proportion living below the poverty line, in count model. The zero-inflation model included the intercept only.

19 risk was 1.13 (95% CI, 0.97–1.32; P = .11). In other words, for this risk period, the estimated odds that a county had a particular risk level or higher, relative to those odds for a county where the voter turnout was about 1 percentage point lower was 1.13, consistent with a substantial positive association. Alternatively, we adjusted for additional covariates (proportions with college education, with a high school education, who smoke and who are obese) using random-effects, zero-inflated, negative binomial models and obtained RRs consistent with those in Tables 7 and 8. (Some supplemental analyses, however, involved models with relatively few observations per parameter, so we view them as secondary; Online Supplement).

Discussion

Our results differ by state. For Michigan, to a lesser extent for Missouri, and not at all for Mississippi, they suggest that counties with a larger voter turnout had higher COVID-19 risks over an approximate 1- to 2-week period beginning a few days after the voting. Our negative control exposure analyses do not permit a definite conclusion. Initial analyses identified positive associations of risk with the negative control, which could indicate confounding, but the associations disappeared or reversed with the exclusion of one influential county (Online Supplement). The results are compatible with either some residual confounding or with circumstantial events (cases) in a single county and no residual confounding.

It is important to consider possible reasons that might explain the heterogeneity of the association. One possibility is that Mississippi, where no association was seen, and Missouri, where a weak or minimal association was seen, had few infectious cases on March 10. If so, we would expect to see little or no association with voting in these states. This hypothesis is plausible because the number of cases in the JHU-CSSE data was five-fold to 10- or more fold greater in Michigan than in the other two states during the first weeks after the election. However, we cannot dismiss other possibilities such as greater confounding in Michigan, use of overall voter turnout as a surrogate for in-person voting in Mississippi and Missouri, differences in social distancing, or the role of chance. Our analyses for Michigan suggest a substantial increase in risk may be associated with higher voter turnout for part of the 2-week risk period after the election, perhaps as much as 20% (median RR, for the main outcomes in Table 8, is 1.2). We view this estimate as...
very approximate because of fairly wide CIs, more moderate RRs in some models (Online Supplement), and its model dependency. Our results also suggest no increase in Mississippi and little, if any, increase in Missouri.

The increase in risk we observed in Michigan was predominantly restricted to the 1–2 weeks after March 13, lending strength to our findings. Importantly the increase was small or absent, either when we included additional days or defined outcome using periods starting in April (alternative outcome analyses). Although the chain of transmission attributable to the voting would be expected to continue after the early postvote period, the effect of voting should be less apparent because its contribution relative to other factors influencing community spread should diminish. We expect the increases would become more diffuse and less county-specific with increasing time, a pattern compatible with our observations. In analogy with negative control outcomes [29], the weak to null association with the alternative outcomes provides support for the interpretation that residual confounding was not important.

An important limitation of our study is the absence of individual-level data and contact tracing. However, our purpose was not to assess the association between an individual’s participation in voting and COVID-19 risk. Rather, we sought to assess population-level differences in COVID risks that can be attributed to different extents of population participation in a single event. We acknowledge that without individual-level data, we cannot adjust for patterns of risk within counties that depend on the joint distribution of covariates or for unknown or unmeasured confounders that differ across counties [32,33].

A second limitation is that voter turnout imperfectly measures the proportion of the population going to the polls and our direct measure of in-person voting is only in Michigan. (However, this difference in the type of information may have had a modest impact: when we used overall voter turnout in place of in-person turnout in Michigan, the RRs for the main outcomes were roughly 25% closer to the null, but still meaningfully elevated; data not shown.) A further limitation is that particularly for states with the alternative outcomes in Michigan and the pattern of limited by the number of observations (counties).

These limitations notwithstanding, the weak to absent associations with the alternative outcomes in Michigan and the pattern of effects, with greater increases, observed soon after the election followed by smaller increases, as the intervals of study expanded provide some evidence, although indirect, that residual confounding or ecologic bias may not be important threats to validity of the observed results. In Missouri, the lack of a consistent pattern and an association of turnout with an alternative outcome in supplemental analyses (Online Supplement) suggest that the observed associations there may reflect chance, residual confounding, or some other phenomenon driving the epidemic.

In summary, we reiterate concerns noted in an open letter from multiple public health officials to the U.S. Senate and House of Representatives [34] that going to the polls in an election can be associated with increased risk of SARS-CoV-2 transmission. Although our study of the March 10 primary elections had limitations, the results are consistent with the concern that higher in-person voter turnout may have led to increases in local risk of infection in at least one state. Depending on the situation as the next vote nears, voters may wish to consider taking advantage of absentee ballots or other available voting options.

Acknowledgments

Process by which someone else could obtain the data and computing code: publicly available (Table 1).
[27] Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. Lancet 2020;395:1054–62.

[28] CDC. Previous Covid-19 case data. 2020; webpage indicated changes in how counts are listed. https://www.cdc.gov/coronavirus/2019-ncov/cases-updates/previouscases.html. [Accessed 18 April 2020].

[29] Lipsitch M, Tchetgen ET, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. Epidemiology 2010;21(3):383.

[30] Flanders WD, Klein M, Strickland M, Darrow L, Sarnat S, Sarnat J, et al. A method for detection of residual confounding and other violations of model assumptions. Epidemiology 2009;20(6):544–5.

[31] Flanders WD, Klein M, Darrow LA, Strickland MJ, Sarnat SE, Sarnat JA, et al. A method for detection of residual confounding in time-series and other observational studies. Epidemiology 2011;22(1):59.

[32] Greenland S, Robins J. Invited commentary: ecologic studies—biases, misconceptions, and counterexamples. Am J Epidemiol 1994;139(8):747–60.

[33] Rothman KJ, Greenland S, Lash TL. Modern epidemiology. Philadelphia, PA: Lippincott Williams & Wilkins; 2008.

[34] Open letter to the US Senate and house of representatives. 2020. https://cdn.americanprogress.org/content/uploads/2020/05/05061221/21DemocracyTeam_finalmailvotingandcovid19.pdf. [Accessed 6 May 2020].