A Review of the Article:

Guerra and Guerra (2021), "Mask mandate and use efficacy for COVID-19 containment in U.S. States," published *by International Research Journal of Public Health*. DOI: 10.28933/irjph-2021-08-1005

Authors:
Natalie C. DuPré[1,†], ScD.; Seyed M. Karimi[2,3,†], PhD; W. Paul McKinney[4], MD; Riten Mitra[5], PhD; Giang Vu[2], PhD; Shaminul Shakib[2], MPH; Bert B. Little[2], PhD

Author Affiliations:

1. Department of Epidemiology and Population Health, School of Public Health and Information Sciences, University of Louisville, Louisville, KY, U.S.A.
2. Department of Health Management and Systems Sciences, School of Public Health and Information Sciences, University of Louisville, Louisville, KY, U.S.A.
3. Louisville Metro Department of Public Health and Wellness, Louisville, KY, U.S.A.
4. Department of Health Promotion and Behavioral Sciences, School of Public Health and Information Sciences, University of Louisville, Louisville, KY, U.S.A.
5. Department of Bioinformatics and Biostatistics, School of Public Health and Information Sciences, University of Louisville, Louisville, KY, U.S.A.

† These Authors had an equal contribution to the review.

15 September 2021

A recent article by Guerra and Guerra (2021) attempted to measure the effect of U.S. states' mask mandates on the growth rate of COVID-19 cases at the state level. They reported no "association between mask mandates or use and reduced COVID-19 spread in U.S. states." This study was frequently downloaded and debated in the State of Kentucky and at the national level. For example, the May 2021 version of the manuscript was commented on 179 times and recommended by 20 readers on the medRxiv by 14 September 2021. By the date, its August 2021 version was commented on 170 times and recommended by medRxiv 55 readers. To address the confusion created by this publication, we extensively reviewed it from the perspective of experts in public health research.

We found this article poorly structured with a sketchy and biased description of the background and a seriously flawed methodology, and improper analysis hence a wrong conclusion with vital consequences.

Review of Background and Literature

Guerra and Guerra (2021) provided an imbalanced review of the related literature and failed to mention numerous rigorous studies that observed substantial protection from viral transmissions with the use of face masks, including SARS-CoV-2. The authors did not discuss the results of studies that tackled the same question finding that state mask mandates were associated with a reduction in COVID-19 cases and deaths. Specifically, the effect of U.S. states' mask mandates on the growth of COVID-19-related rates was not addressed as others have (e.g., Wei and Wehby 2020, Joo et al. 2021, Guy et al. 2021, and Chernozhukov et al. 2020). In **Appendix 1**, we have provided a review of this literature.

Importantly, Guerra and Guerra (2021) misrepresented the results of some studies. For example, Guerra and Guerra (2021) wrote: "A randomized controlled trial [RCT] of Danish volunteers found no protective benefit of medical masks against COVID-19 infection [8]." Reference #8 (Bundgaard et al. 2020, *Annals of Internal Medicine*) is a randomized controlled trial comparing a group that received a mask recommendation to a group that did not receive a mask recommendation. The study's results do, in fact, suggest a benefit of a mask recommendation relative to the no recommendation group (Odds Ratio=0.82, 95% Confidence Interval 0.54−1.23), though the difference was not statistically significant likely due to the small sample size, potential contamination of the control group, and/or because the intervention group had a higher proportion who had riskier jobs than the control group (8.7% were shop employees or cashiers in the intervention group, whereas only 7.3% in the control group). Also, Bungaard et al.

(2020) did not report the proportion of mask-wearing in the control group who were randomized not to receive a mask recommendation. The effectiveness of mask recommendations would have been less beneficial by construction if those in the control group were wearing masks, making them more similar to the intervention group.

Also, Guerra and Guerra (2021) wrote: "In RCTs before COVID-19, viral infections were not lower in Vietnamese clinicians who wore cloth or medical masks than in the control arm [9]". Reference #9 (MacIntyre et al. 2015, *BMJ Open*) was intended to measure the efficacy of cloth masks and medical masks relative to the standard practice in Vietnam. Contrary to Guerra and Guerra's (2021) claim, MacIntyre et al. (2015) found that the risk of infection was the lowest in the medical mask arm of the trial (0.2% for influenza-like illness), low in the group with standard mask-wearing practice (0.7%), and the highest in the cloth mask group arm at 2.3% (MacIntyre et al. 2015, Figure 2 and Table 2). MacIntyre et al. (2015) indicated that wearing a medical mask was more common in the standard practice arm than the cloth mask arm of the study.

Review of Methodology

Guerra and Guerra (2021) used an overly simplistic approach, although described in a complicated manner, to analyze the effect of U.S. states' mask mandates on state-level case growth rates. Their approach has two main parts: (1) finding states' minimum and maximum case growth rates from July 2020 to March 2021, and (2) plotting the maxima and minima for five groups of states based on the timing of mask mandates. Using plots and overly simplistic statistics, they made conclusions about the effectiveness of mask mandates.

This approach has serious methodological flaws. One major flaw is not accounting for confounding factors such as other state policies and events that could significantly influence the effectiveness of mask mandates (which were typically implemented in summer 2020) on the maximum growth rate of COVID-19 cases (took place in late November or early December 2020). The influence of confounding factors on mask mandates' effectiveness would be smaller if the effectiveness was measured not long after the mandates. Therefore, Guerra and Guerra (2021) reported a beneficial correlation between the mandates and the minimum case growth rates (Figures 1C-1E), which took place in September 2020. Nonetheless, they de-emphasized this finding and based their conclusion on the more severely confounded finding on maximum cases growth rates. An additional major flaw, which is described in more detail in **Appendix 2**, is that Guerra and Guerra (2021) are comparing growth rates across five small groups (Group 1: states with a mask mandate in April 18-May 16; Group 2: mask mandate in

May 29-July 3; Group 3: July 8-July 27; Group 4: August 1-December 9; Group 5: no mask mandates) with about 9 to 12 states in each group. With such small sample sizes in each group, it is unlikely to statistically significant differences would be detectable across the groups. Essentially, Guerra and Guerra (2021) are comparing small groups, ensuring a lack of statistically significant differences across the five groups to support their viewpoint.

In addition, Guerra and Guerra's (2021) empirical approach is highly susceptible to the ecologic fallacy. They did not examine if the COVID-19 cases were using masks or not, so they cannot make any individual-level conclusions. Yet, the ramifications of this ecological article are that mask mandates are ineffective. At the same time, the study shows that states with mask mandates had higher usage of masks (Figure 1I), while the preponderous of literature on the effectiveness of masks indicates reduced viral transmissions.

Differential outcome measurement bias is another important flaw of Guerra and Guerra's (2021) study. They compare case growth in states between the modeled effective mask mandate date to March 2021, but the start date differs for each state. The issue is that, by the authors' definition, states with earlier mandates will have more *total* cases per 100,000 between an earlier date until March 2021. This is differential outcome measurement bias when case growth is being defined differently based on the mask mandate status.

Other methodological problems of Guerra and Guerra's (2021) study are simultaneity bias, omitted variable bias, selection bias, statistical inference at the state-level with small sample sizes, lack of information on key data, use of questionable data sources, incorrect statistical terminology, unconventional methodologies for adjusting for population density, unclear description of methods and results, and contradictory results. We have explained the implications of this problem for the results of this study in **Appendix 2**.

Appendix 1:

A Review of the Literature on Effect of Mask-Wearing on Reducing the Risk of Respiratory Infections

Our review of the literature on the effect of mask-wearing on reducing the risk of respiratory infections, including SARS-CoV-2, showed the high efficacy of mask-wearing. We specifically focused on the methodology used in the reviewed studies and put them in five categories: (1) systematic reviews, (2) laboratory studies on SARS-CoV-2, (3) mask recommendation randomized controlled trials during COVID-19, (4) cohort and case-control studies, and (5) assessments of the effectiveness of mask mandates.

Systematic Reviews

Because of ethical and logistical reasons, evidence on the efficacy of facial masks is usually from observational studies. A systematic review of such studies (30 studies, 27 were peer-reviewed), published in *The Lancet* in June 2020, concluded that mask-wearing can reduce the risk of infection from respiratory viruses by 85% (95% CI: 66%−93%) (Chu et al. 2020). In January 2021, another article, published in *Proceeding of the National Academy of Science (PNAS)*, conducted an updated review and came to a similar conclusion: facial mask-wearing has high efficacy (Howard et al. 2021). The latter article also showed that mask-wearing efficacy significantly increased when compliance to mask-wearing was high.

A systematic review focused on SARS-CoV-2, published on August 2021 in the *EClinicalMedicine*, was in line with the finding of systematic reviews of other respiratory infections (Ford et al. 2021).

Laboratory Studies on SARS-CoV-2

Laboratory studies were the earliest to assess the effect of mask-wearing on the transmission of SARS-CoV-2. Researchers typically measured the distances that droplets containing SARS-Cov-2 transport with and without a facial mask in a laboratory environment (e.g., Abkarian et al. 2020 *PNAS*, Bahl et al. 2020 *Thorax*, and Fischer et al. 2020 *Science,* Verma et al. 2020 *Physics of Fluids,* Ueki et al. 2020 *mSphere*, Lindsley et al. 2021 *Aerosol Science and Technology*). The findings of these studies strongly support mask-wearing.

Mask Recommendation Randomized Controlled Trials during COVID-19

Two randomized controlled trials (RCTs) in the mask-wearing/COVID-19 transmission literature were located. In the trials, mask-wearing was promoted in a randomly selected group of individuals with known compliance rates while not recommended to a control group.

The results of one study, conducted in Denmark in April and May 2020 and published in the *Annals of Internal Medicine* in November 2020, suggested a benefit of a mask recommendation relative to the no recommendation group, though it is not statistically significant (OR=0.82, 95% CI 0.54−1.23) likely due to the small number of COVID-19 cases in each group, n=42 in the face mask group and 53 in the control group (Bundgaard et al. 2020). This study did not assess the proportion of mask-wearing in the group who were randomized not to receive a mask recommendation, who may have worn masks as frequently as the treatment group. Therefore, it is likely that the estimated difference in SARS-CoV-2 infection between the two groups is underestimated. In addition, the group randomized to receive the mask recommendations had a higher prevalence of workers in high-risk settings (e.g., shop workers and clerks) than the group that did not receive the mask recommendation.

In a much larger RCT study in Bangladesh (involved 341,830 adults), published as an *National Bureau of Economic Research* (*NBER*) working paper in April 2021, researchers examined the effectiveness of an array of strategies to promote mask-wearing and found periodic monitoring as the most effective strategy, resulting in the tripling of mask use rate in comparison to the control group (Abaluck et al. 2021). The effect on the transmission of SARS-CoV-2 is not yet reported by this study's authors.

Cohort and Case-Control Studies

Three examples of these types of studies in the COVID-19 era include (1) Wang et al. (2020), conducted in China from 28 February to 27 March 2020, published in the *BMJ Global Health* in May 2020, (2) Payne et al. (2020), used data from U.S. Navy service members on the USS Theodore Roosevelt in April 2020, published in the *Morbidity and Mortality Weekly Report (MMWR)* in June 2020, and (3) Doung-Ngren et al. (2020), conducted in Thailand from 30 April to 27 May 2020, published in the *Emerging Infectious Diseases* in November 2020. All three studies found a significant effect of mask-wearing on decreasing the risk of SARS-CoV-2 infection.

Assessments the Effectiveness of Mask Mandates

Methodologically sound observational studies on the effect of U.S. states' mask mandates show the effectiveness of the mandates on growth rates of COVID-19 cases, hospitalizations, and deaths at the county level. The studies attempted to isolate the effect of mask mandates from other concurrent factors and measured effects that were realized from 5 to 100 days after the implementation of the mandates.

For example, one study, published in June 2020 in the journal *Health Affairs*, analyzed COVID-19 case growth rate from 31 March to 22 May 2020 and estimated 0.9 to 2.0 percentage point decreases in the rate in 5-day intervals after a state's mask mandate in comparison to its pre-mandate rate and the rate in non-mandate states (Wei and Wehby 2020).

Another study, published in the Center for Disease Control and Preventions's (CDC) *MMWR* journal in February 2021, analyzed the COVID-19 hospitalization growth rate from 21 March to 17 October 2020. Hospitalizations decreased by 5.6 percentage points among adults during the first two weeks and from three weeks after the mandate in the implementing states versus non-implementing states (Joo et al. 2021).

Another study, published in the CDC's *MMWR* journal in March 2021, analyzed 1 March to 31 December 2020 data. They measured a 0.7 to 1.9 percentage point decline in death growth rates in 20-day intervals after a mandate in a typical implementing state versus a non-implementing state up to 100 days, 1.1 to 1.8 percentage point declines in death growth rates (Guy et al. 2021).

Other carefully designed studies on the effect of mask mandates in the U.S. confirm the results of these three studies (e.g., Chernozhukov et al. 2020 *Economic Journal*). Studies of data from other countries have found similar results. Two prominent examples are a study using data from Australia, published in *PLOS ONE* (Scott et al. 2021), and a study using data from Canada, published in *the Journal of Health Economics* (Karaivanov et al. 2021). Both were published in July 2021.

Appendix 2:

A Detailed Review of Guerra and Guerra's (2021) Methodology and Result

The core of Guerra and Guerra's (2021) analysis can be summarized in two parts:

(1) Finding minimum and maximum case growth rates for the U.S. states from 23 July 2020 to 6 March 2021 (49 minima and 49 maxima were founded, excluding Hawaii). The minima usually took place in September 2020, the maxima in late November or early December 2020 (Guerra and Guerra's 2021, Figure 1, A-B).

(2) Plotting minimum and maximum case growth rates for U.S. states, grouped to five based on the timing of mask mandate. Using plots and basic statistics, they made statements about the correlations between the state-level minimum and maximum case growth rates and mask mandates. Most mask mandates were put in place before the end of summer 2020.

This overly simplistic approach, however, is seriously flawed. In the following, we reviewed the flaws.

Influence of Confounding Factors

Understandably, Guerra and Guerra's (2021) visualize no correlation between mask mandate and the maximum case growth rate because maximum growth rates took place months after the mask mandates were implemented (Guerra and Guerra 2021, Figures 1F-1H). Many other things happened in between, and the effect of a mask mandate can be confounded by the effect of other interventions, events, policies, and adherence to the mandates. Again, understandably, they observe a correlation between mask mandate and the minimum case growth rate because the timing of minimum growth rates and mask mandates were temporally close (Guerra and Guerra 2021, Figures 1C-1E). Nonetheless, the visual correlation between the minimum case growth rate and mask mandate can also be spurious, and the lack of statistical significance is nearly guaranteed due to their small sample sizes.

Ecologic Fallacy

Another major limitation of Guerra and Guerra's (2021) study is its high susceptibility to the ecologic fallacy. The authors of this paper compared COVID-19 case growth in states with early versus late mask mandates to states without a mask mandate. Nowhere in this ecological study is it stated whether the COVID-19 cases were using masks or not, so we cannot make any individual-level conclusions. Yet, the ramifications of Guerra and Guerra's (2021) ecological article are that mask mandates are ineffective. At the same time, the study showed that states with mask mandates had higher usage of masks. The preponderous of peer-reviewed literature confirms the effectiveness of masks at reducing viral transmissions, albeit of other viruses similar to SARS-CoV-2.

Simultaneity Bias

A mask mandate may follow increasing COVID-19 trends in a state. In other words, states with worse COVID-19 numbers or those that experienced it before other states might be more likely to have a mask mandate earlier. This will result in underestimating the true effect of a mask mandate on COVID-19 transmission. The authors needed to check this possibility and carefully account for it, but they did not.

Omitted Variable Bias

Even if the simultaneity bias is carefully addressed, the authors needed to adjust for sociodemographic factors. This study uses simple statistical tests and does not account for sociodemographic and socioeconomic factors, except for population density in some tests. There are other differences in addition to the population size and density (for example, age structure, labor market and economic structure) in states with a mandate versus no mandate that could mask an association when there truly is an effect.

Differential Outcome Measurement Bias

Figures 2B-2D of Guerra and Guerra's (2021) study compare case growth in states between the modeled effective mask mandate date to March 2021, but the start date differs for each state. The issue is that, by the authors' definition, states with earlier mandates will have more *total* cases per 100,000 between an earlier date until March 2021. This is differential outcome measurement bias when case growth is being defined differently based on the mask mandate status.

Selection Bias

It is difficult to verify which analyses excluded the Northeast states. Guerra and Guerra (2021) wrote: "For the Summer wave, Northeast states were excluded because they deviated from other states with respect to total cases and growth covariation." However, it seems like the Northeast states are likely excluded from the main results in Figures 1C-1H that evaluate July 2020-March 2021 case growth. Excluding states with the lowest case rates, which were adopters of the mask mandates, the authors removed the group that showed a benefit of a mask mandate. This is called selection bias and, in this case, would lead to observing no effect of a mask mandated, even if they were a protective effect of a mask mandate.

Inference at the State-Level with Small Sample Sizes

The authors rely on statistical significance to say that states are not different. However, states were grouped in small sample sizes (9−12 states in each group). Therefore, it will be very difficult to detect statistical differences between small sample number groups.

The strongest conclusions are coming from the article's Figures 1C-1H that define COVID-19 growth starting from the end of the Summer wave (23 July 2020) to the height of the Fall-Winter wave (1 March 2021). However, no table presents p-values, sample sizes, or average differences in case growth between the groups. Most of the results of their analyses are not presented. It is poor science to leave readers' conclusions to rely upon the authors' subjective statements and misuse statistical significance in the context of small sample sizes. Moreover, lay individuals do not detect these missteps in science methodology and believe the ill-informed analyses. For example, Guerra and Guerra "adjust" for population density (Figures 1E and 1H) and report differences between states as statistically significant. No differences are detected without population adjustment. With as few as 9 states in each group, there is little statistical power to detect such differences. Nonetheless, Figures 1E and 1H are visually more in line with the notion that states with early mask mandate adoption had lower case rate growths than states with later adoption or no adoption of mask mandates.

Lack of Information on Key data and Use of Questionable Data Sources

- Guerra and Guerra (2021) define the timing of mask mandate as "effective dates," which is their main exposure of interest in this paper, "were obtained from US state executive and health departments and press releases [available upon request]."

- Mask usage data from the Institute for Health Metrics and Evaluation's (IHME) secondary data sources and surveys such as Facebook and Kaiser that are not validated and likely representative of the state population, nor is it known if they measured mask usage similarly across U.S. states. Therefore, figures on mask usage (Guerra and Guerra 2021, Figures 3 and 4) are highly subjective and should be totally ignored because the sources of mask usage data are questionable, and there is no mention of whether this data on mask usage is valid.

- Testing for SARS-CoV-2 was not uniform across states, and the study did not consider COVID-19 deaths or hospitalizations that are a better metric of COVID-19 across states than infection cases.

<u>Incorrect Statistical Terminology</u>

Guerra and Guerra incorrectly use the terms "normalized cases," and mask mandate "quintiles."

- The term "normalized" cases is incorrect; normalization is the process of transforming the data to a Gaussian or normal distribution with a mean of 0 and a standard deviation of σ. Instead, the term that is used in biostatistics and epidemiology is a case rate, or the number of new cases per 100,000 people, which is what Guerra and Guerra appear to use though it is not always clearly stated throughout the manuscript.

- Quintiles are when the dataset is divided into five equal parts (20% in the lowest quintile, 20%, 20%, 20%, and 20% of the data in the highest quintile). Guerra and Guerra say: "We assigned US states to one of five quintiles based on when mandates went into effect [effective dates]: 18 April-16 May 2020 [Q1], 29 May-3 July 2020 [Q2], 8 July-27 July 2020 [Q3], 1 Aug-9 Dec 2020 [Q4], or no statewide mandate as of 6 March 2021 [Q5]."

  However, these are really just five categories based on their definitions of the timing of mask mandates, and in Figure 2A, the dates of the effective mask mandate do NOT match their categories; for example, New Hampshire and Vermont have a "mandate effective date" of 4/27 and 5/03 but are in Q4 that should be 1 Aug-9 to Dec 2020. The net effect is it is very difficult to find the number of states in each of the five categories because the authors do not ever state their sample sizes.

Unconventional Methodologies for Adjusting for Population Density

- The analysis could have been done in a much less complicated fashion using Poisson regression or log-linear regression as done in Guerra and Guerra's References #15 and #16 (Guy et al. 2021 and Wei and Wehby 2020, respectively).

- The authors "normalize" their fold growth outcome for each state by urban and rural population density; however, outcome data is not available for urban and rural areas.

Unclear Description of Methods and Results and Contradictory Results

The methods and results are unclear, contradictory, and difficult to follow (i.e., poor organization of the methods and results), and the strong conclusion statements from the results are ungrounded and cherry-picked.

- Guerra and Guerra's (2021) Figure 1J appears to represent different time periods of total cases than the rest of the Figure 1 figures, but the authors did not clearly present these data.

- For example, the authors stated, "Adjusting for population density, Fold maxima were 2.8-fold higher in MQ4 than MQ2 [p=0.01], but MQ1, 3, and 5 were indistinguishable [Fig. 1H], suggesting mask mandate duration was not associated with lower maximum growth." The abbreviation MQ, not defined in the text, seems to stand for mandate quintiles, but what do the authors mean by indistinguishable? There is no presentation of statistics on average growth rates to support this assessment. The lack of finding statistically significant differences is more likely due to the small sample sizes, and visually, the groups do look different after "adjusting for population density" in their unconventional methodology.

- This is a clearly contradictory statement after the results show borderline statistically significant results: "total cases on 1 March 2021 were MQ-independent [p<0.07; Fig. 1J]. Direct MQ1 vs. MQ5 comparison by t test uncovered a small [1.2-fold] and non-significant [p=0.078] difference in total cases. Taken together, these findings suggest that US state mask mandates were not associated with slower spread of COVID-19."