



The Authors' Reply: Need for a multi-pronged population-level strategy to manage SARS-CoV-2 infection

S. V. Subramanian^{1,2,3} · Akhil Kumar²

Received: 12 November 2021 / Accepted: 2 December 2021 / Published online: 24 December 2021
© Springer Nature B.V. 2021

We thank Mulot, et al. [1], Coleman, et al. [2], and Gianicolo, et al. [3] for their responses to our research letter on the lack of a meaningful association between incidence in confirmed COVID-19 cases and levels of full vaccination across 68 countries and 2947 counties in United States (US).

The goal of our analysis was not to discredit the idea that vaccination is highly effective at preventing severe disease, hospitalization, or death at the individual level. It was instead to investigate the increase in confirmed cases that several populations were experiencing due to the Delta variant, as well as known concerns over waning immunity from vaccinations [4]. Within this specific context, and from what the data suggested, we concluded that even as we encourage people to get vaccinated, consideration of other known non-pharmacological interventions was equally necessary [5].

In this rejoinder, we respond to the concerns raised by the commentators and endeavor to correct any misinterpretations of our work. In order to uphold the spirit of a respectful and scientific exchange, we restrict our responses to the empirical comments made by the commentators with regards to our research letter.

On the data

The data source we used for the US county-level analysis was the public data set that the White House Task Force and Centers for Disease Control and Prevention (CDC) make available on a frequent basis [6], referred to as the

COVID-19 Community Profile Report. We chose this data set because it is being used by the White House Task Force and CDC on various aspects of monitoring, managing and mitigating the COVID-19 pandemic [7]. As a way of aligning to this policy effort, as well as respecting the expertise of the agencies involved, we gave salience to the way the variables are reported in the data set.

We recognize that occasionally there may be discrepancies between the data provided by the White House Task Force/CDC and websites of the respective US counties, but it would have been unscientific for us to pick and drop different values from the data set. We could not rationally discount the very data set that is being used to shape and manage the COVID-19 response in the US [8].

The only exclusion we made was to drop counties in the data set that did not provide the percentage of population fully vaccinated.

At the country level, we used the data set provided by *Our World In Data* [9], collated by the respectable academic team at the University of Oxford and Global Change Data Lab. We chose this source because of their longstanding experience in collating and providing data on a large range of global problems with detailed notes on the original sources [10]. As outlined in the original research letter [5], we developed an inclusion/exclusion criterion to include countries with most recent data updates, and hence the analysis was restricted to countries for whom “*the last update of data was within 3 days prior to or on September 3, 2021*” [5]. Following this criterion, countries with older data on vaccination and/or case metrics were not included. For example, France and the United Kingdom were excluded since they last updated their vaccine data on August 31, 2021 at the time the original research letter was being written.

✉ S. V. Subramanian
svsubram@hsph.harvard.edu

¹ Harvard Center for Population and Development Studies, Cambridge, MA, USA

² Center for Geographic Analysis, Harvard University, Cambridge, MA, USA

³ Department of Social and Behavioral Science, Harvard T.H. Chan School of Public Health, Boston, MA, USA

On considering cases as the variable of interest

We believe an analysis of cases in aggregate can be meaningful for assessing whether vaccination can be considered as a sole prevention measure in stopping infection and transmission at the population level. Further, the use of aggregate confirmed case data remains the most used variable both in the public and policy deliberations and narratives [7]. In choosing this widely used metric to understand COVID-19, we followed what has been the established academic and policy practice of normalizing the case variable as a “per persons” metric.

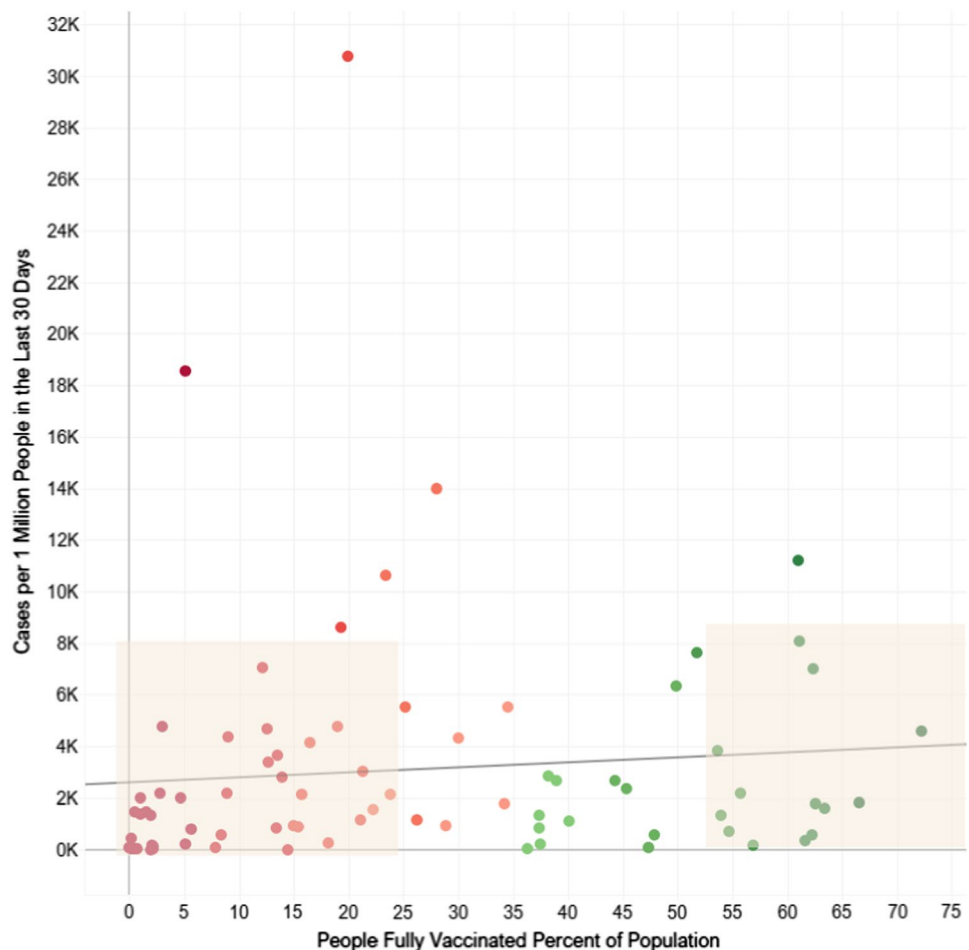
On the rationale for the “weekly” window for defining incident cases

In deciding the precise time-window to define “incidence” we relied on what was provided in the data set. The White House Task Force/CDC data readily provides the 7-day

window for cases by counties for various dates. Since a team of interdisciplinary experts oversee this data collation and to keep our analysis aligned with the data set being actively used for policy formulation, we choose to keep our analysis closest to what was being made available officially.

In our initial submission, we had presented a 30-day window for the country level analysis, where we had a choice to define the time window. However, during the external peer review process, it was suggested that we change the country-level analysis time window to a 7-day period in order to truly capture “incidence”. We include here the analogous scatter plot for the 30-day period for the countries from our initial submission, albeit for a different date (August 10, 2021) (Fig. 1). As seen in the figure, our overall interpretation of no discernible association between percentage of people fully vaccinated and cases per 1 million people in the last 30 days appears to hold, even when there are 11 more countries and a larger time frame of 30 days.

Fig. 1 People fully vaccinated as percent of population and cases per 1 million people in the last 30 days across 79 countries as of August 10, 2021



On miscellaneous aspects of the analysis

First, our assessment of percentage of counties experiencing an increase over the two time windows in the original research letter did not consider the “size of the increase”. We examined the percent change in cases over the two time periods and the percentage of people who are fully vaccinated across US counties (Fig. 2). As is evident, the overall inference drawn in the research letter remains unaltered.

Second, we wish to reiterate that in the original letter we already present a sensitivity analysis, in the associated Supplementary File, where we considered a one-month lag on the independent variable “percentage of population fully vaccinated” and confirmed the robustness of our findings for both countries and US counties [5].

Third, in our research letter, we also included an interactive data dashboard (<https://tiny.cc/USDashboard>) that allows users to visualize the patterning of cases per 100,000 people in the last 7 days and percentage of population fully vaccinated (in categories) along with other

data metrics since April 12, 2021, with automatic updates as new data is released.

Fourth, we elected to use boxplots across various categories of percentage of population fully vaccinated because it was the most appropriate way to describe the underlying data, which is the primary goal of any statistical analysis. Specifically, in the case of the two variables we focused on, the variation between counties *within* a category of fully vaccinated is readily apparent—especially in the US—regardless of the average for any category.

We thank Mulot et al. for analyzing the two variables from the same data source we considered in the analysis using more formal statistical models (Supplementary Analysis of their comment) [1] that analyze the relationship using a continuous scale of fully vaccinated. Mulot et al. selectively interpret the higher end of the distribution of fully vaccinated to suggest a negative correlation with “cases per 100,000 in the last 7 days” [1]. However, the variation between counties at higher levels of vaccination is evident in their figure, reinforcing our interpretation that a county could be high or low in cases at different levels of

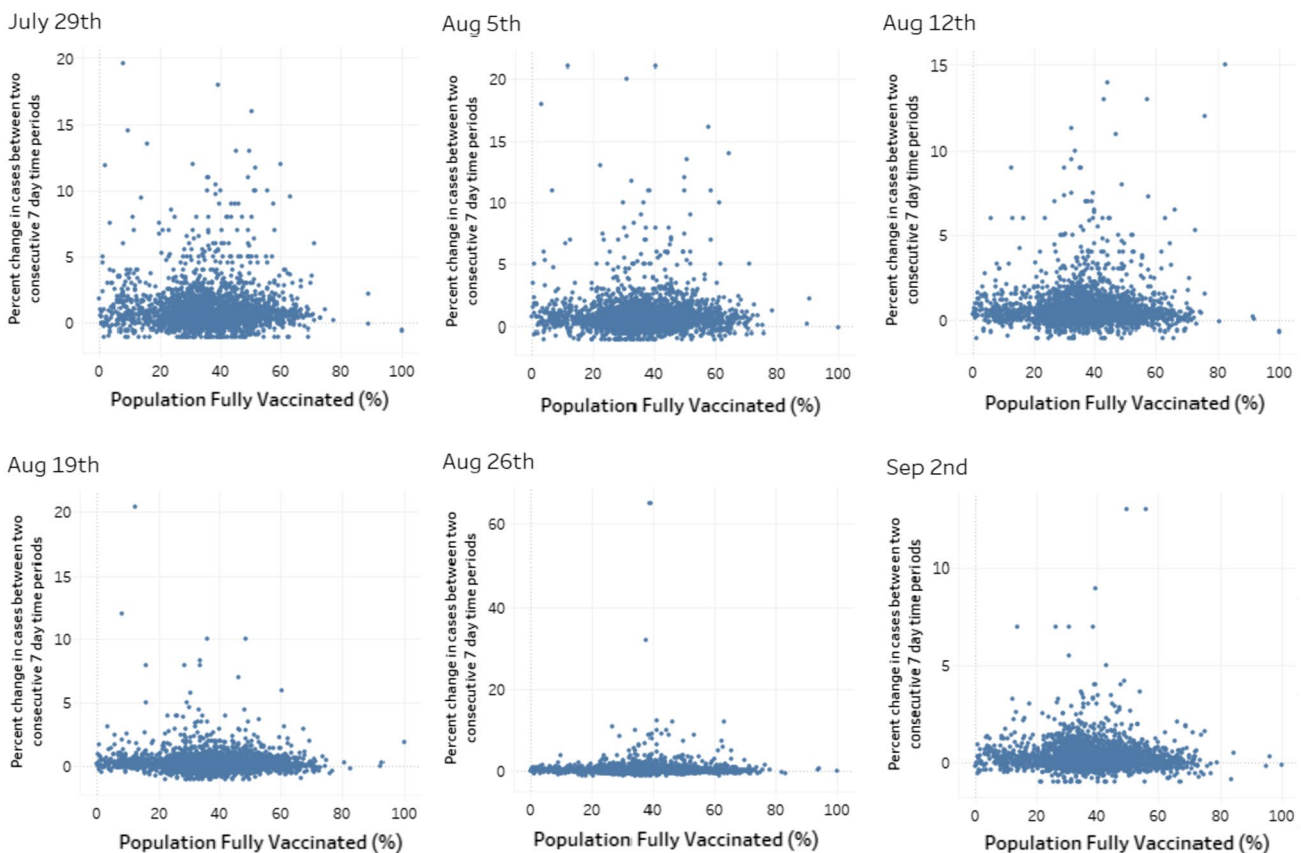


Fig. 2 Percent change in cases between two consecutive 7-day time periods and percentage of population fully vaccinated for 6 different weekly time periods going back until the closest date to July 27, 2021

since that is when the CDC added more restrictive guidelines (*i.e.*, masking measures) for fully vaccinated people citing the Delta variant

vaccination, and the appropriateness of boxplots to characterize the data.

It is precisely because of this finding we concluded that there should be a consideration of other known non-pharmacological interventions in addition to vaccines. For instance, in a recently published simulation study of a university campus, it was shown that surveillance testing, and isolation of positive cases are important mitigation strategies, even if 100 percent of the students are vaccinated [11].

Finally, we concur with the concern around observed and unobserved confounding variables in any observational data analysis. This is especially true for the country-level analysis. While we acknowledge this particular limitation in our original research letter, space constraints inhibited us from elaborating in greater depth. It is precisely because of our recognition of a whole variety of country-level differences that we deliberately restrained from overinterpreting the statistically positive counterintuitive association observed across countries. Instead, we leaned towards interpreting an association that is descriptively self-evident.

On the target of inference: population vs individual

Our analysis and interpretation does not commit ecological fallacy, *i.e.*, interpreting results observed at a group-level and transferring them incorrectly to the individual-level; a subject on which one of us (SVS) has extensively engaged [12, 13]. It is not unusual in epidemiological research (and especially in its public communication) to conflate the two targets of scientific inference: populations vs individuals [14]. Björk et al., touch upon this in their commentary as well [15]. In their comment they present an ecological analysis at the state level and conclude a positive association between percent fully vaccinated and cases per 100,000 people in the last 7 days. It may be noted that the population level units of states and counties are not interchangeable, a point well demonstrated by W.S. Robinson more than 70 years ago [16]. Furthermore, even the state level scatterplot appears to show substantial heterogeneity especially below 55% fully vaccinated, making the use of boxplots as a better description of the underlying data.

In short, whether a group-level data analysis, individual data analysis, or multilevel data analysis is appropriate depends on the question of interest. The inferential target of our analysis was *population*. The transmissibility of the virus makes it appropriate for a group-level (*i.e.*, population level) analysis to assess changes in cases. Indeed, if the goal were to assess the individual risk of infection or

hospitalization or mortality, an analysis of individual-level data would be the appropriate strategy.

Concluding remarks

The concerns raised in the comments does not necessitate altering our original inferences and conclusions [5]. Nonetheless, we hope that future research will continue to utilize a data-oriented approach to inform this important question of the association between cases and levels of fully vaccinated.

Over the course of this pandemic, if we have learned one thing, it is that the virus is always ahead of us. Using all the tools at our disposal—handwashing, masks, physical distancing, proper ventilation, testing as well as vaccines—will give us the greatest protection possible at the individual and population level.

Acknowledgements We wish to acknowledge the responses we have heard on our original research letter outside of the *Journal* platform. We have made a sincere attempt in our rejoinder to consider those that fall under the purview of a scientific discourse.

References

1. Mulot M, Segalas C, Leyrat C, Besancon L. Vaccination rates and COVID-19 cases: a commentary of “Increases in COVID-19 are unrelated to levels of vaccination across 68 countries and 2947 counties in the United States.” *European Journal of Epidemiology*. 2021. <https://doi.org/10.1007/s10654-021-00817-6>.
2. Coleman T, Sarian A, Grad S. Commentary on “Increases in COVID-19 are unrelated to levels of vaccination across 68 countries and 2947 counties in the United States.” *European Journal of Epidemiology*. 2021. <https://doi.org/10.1007/s10654-021-00821-w>.
3. Gianicolo E, Wollschläger D, Blettner M. Comment on „Increases in COVID-19 are unrelated to levels of vaccination across 68 countries and 2947 counties in the United States“. *European Journal of Epidemiology*. <https://doi.org/10.1007/s10654-021-00816-7>.
4. Taylor L. Why scientists worldwide are watching UK covid infections. *Nature*. 2021. <https://doi.org/10.1038/d41586-021-03003-6>.
5. Subramanian SV, Kumar A. Increases in COVID-19 are unrelated to levels of vaccination across 68 countries and 2947 counties in the United States. *European Journal of Epidemiology*. 2021. <https://doi.org/10.1007/s10654-021-00808-7>.
6. White House COVID-19 Team. COVID-19 community profile report. <https://healthdata.gov/Health/COVID-19-Community-Profile-Report/gqxm-d9w9>. Accessed 11 Dec 2021.
7. White House. Press Briefing by White House COVID-19 Response Team and Public Health Officials. 2021. <https://www.whitehouse.gov/briefing-room/press-briefings/2021/05/21/press-briefing-by-white-house-covid-19-response-team-and-public-health-officials-38/>. Accessed 11 Dec 2021.
8. CareSet CDC. CareSet/covid_community_profile_report_faq: The FAQ for the Community Profile Report. GitHub 2021. https://github.com/CareSet/COVID_Community_Profile_Report_FAQ. Accessed 11 Dec 2021.

9. Ritchie H, Ortiz-Ospina E, Beltekian D, Mathieu E, Hasell J, Macdonald B, Giattino C, Appel C, Rod s-Guirao L, Roser M. Coronavirus pandemic (COVID-19). 2020. Published online at OurWorldInData.org. <https://ourworldindata.org/coronavirus>. Accessed 11 Dec 2021.
10. Our World In Data. About. 2021 <https://ourworldindata.org/about>. Accessed 11 Dec 2021.
11. Motta FC, McGoff KA, Deckard A, et al. Assessment of Simulated Surveillance Testing and Quarantine in a SARS-CoV-2–Vaccinated Population of Students on a University Campus. *JAMA Health Forum*. 2021. <https://doi.org/10.1001/jamahealthforum.2021.3035>.
12. Subramanian SV, Jones K, Kaddour A, Krieger N, Revisiting Robinson: The perils of individualistic and ecologic fallacy. *Int J Epidemiol*. 2009;38:342–60. <https://doi.org/10.1093/ije/dyn359>.
13. Subramanian SV, Jones K, Kaddour A, Krieger N, Response: The value of a historically informed multilevel analysis of Robinson's data. *Int J Epidemiol*. 2009;38:370–3. <https://doi.org/10.1093/ije/dyn354>.
14. Subramanian SV, Kim R. Sick countries and sick individuals: reassessing the inferential targets in global health research. *Lancet Glob Health* 2021;6:e726–7. [https://doi.org/10.1016/s2214-109x\(18\)30262-6](https://doi.org/10.1016/s2214-109x(18)30262-6)
15. Bj rk J, Modig K, Kahn F, Ahlbom A. Revival of ecological studies during the COVID-19 pandemic. *European Journal of Epidemiology*. 2021. <https://doi.org/10.1007/s10654-021-00830-9>
16. Robinson, W. S. Ecological Correlations and the Behavior of Individuals. *Am Sociol Rev*. 1950;15:351–7. <https://doi.org/10.2307/2087176>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.