PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	US-county level variation in intersecting individual, household, and community characteristics relevant to COVID-19 and planning an equitable response: A cross-sectional analysis
AUTHORS	Chin, Taylor; Kahn, Rebecca; Li, Ruoran; Chen, Jarvis; Krieger, Nancy; Buckee, Caroline; Balsari, Satchit; Kiang, Mathew

VERSION 1 – REVIEW

REVIEWER	Danielle N. Poole Dartmouth College, United States
REVIEW RETURNED	14-May-2020

GENERAL COMMENTS	Overall review and general recommendation: The paper addressed the important topic of intersecting risk factors in local responses to COVID-19 for effective containment. The authors present five bivariate analyses of risk factors at the county level in the United States. The findings and interpretation have the potential to be of high impact, and the online tool is a great contribution to the COVID-19 response. However, tightening the narrative and motivation would improve this important paper. Therefore, I recommend minor revisions. Below I detail my concerns. Major comments: 1. The Introduction is a bit jumpy – paragraph 1 discusses healthcare resources, poverty, and race, paragraph 2 addresses age, pre-existing conditions, population density, living arrangements, and poverty again. Organizing along the themes of biologic, demographic, and socioeconomic risk factors, recognizing the overlaps, might improve the structure of the Introduction. 2. Why the analysis is done at the county level remains unclear. What decisions are made at the county vs. state level, and why is this important? Social distancing policies are mentioned, but those decisions are typically at the state level. The Results are also described at the state level (e.g., Texas, Georgia, Mississippi) While locally tailoring approaches is logical, motivation and interpretation for this particular level of analysis is missing. 3. The paper starts and ends with a discussion of race (first paragraph of Introduction and Conclusion), but this focus isn't clear throughout the paper where race appears in only one of the bivariate analyses. Important intersections of race with several of the other risk factors could have been examined to strengthen this narrative, and would have been very interesting. However, it seems the purpose of this paper is to present examples of intersecting risk, which is also valuable. Clarification on why these five analyses are presented, instead of other potential analyses, would be helpful.

	Minor comments: 1. Page 5, lines 24-27: Overall, it's unclear why these are the five analyses presented (out of 24 available risk factors). Adding numbers to the separate analyses could help clarify this sentence. 2. Page 6, line 12: Why wasn't the public (i.e. county-level decision makers) involved, especially as the tool is built for them? 3. Page 7, lines 2-4: New Mexico, Arizona, and Colorado appear to have the highest concentration of counties with high poverty and populations over 70 (have the most dark purple shading), but aren't mentioned, whereas Georgia, Texas, and Arkansas are highlighted. 4. Page 10, lines 14-18: What's the standard for bed capacity per 100,000? 5. Page 14, lines 31-35: Useful proposal of modifying these risk factors through financial aid and healthcare. Is there any evidence of financial aid and improved social distancing? 6. There are a few copy-editing issues, i.e. "U.S." versus "US" versus "United States," missing period on page 12, ect.

REVIEWER	Avirup Guha Case Western Reserve University, Cleveland, OH
REVIEW RETURNED	22-May-2020

GENERAL COMMENTS	Thank you bringing out this effort to look at county level data regarding social disparities which exist in America. Perhaps the connection this group is making and conclusions being drawn about the social determinants of health are correct. However, there are major flaws which requires to completely change the approach to make this publishable. I will leave some clues here to help take this to the next level: 1) Always start with references and work which is peer reviewed and published. There is a lot of referencing of data which is from news media which might be biased and unscientific. Speaking of references, please reference introduction, methods and discussion. Results section should talk only about original data. 2) Methods: I see a lot of maps which can be drawn using CDC public website. There is not much of science here. I would however had been interested in a 2*2 figure where in number of positive cases of COVID would have been put together with these variables/social determinant. Merely looking at current maps are thought provoking but not scientific enough. Along the same lines some amount of basic logistic and linear regression modeling would improve ability to draw conclusions. 3) Results: There is too many things going in the results section to draw somewhat homogeneous conclusion. I recommend taking the high impact message and putting it in first and then presenting smaller facts fully in the supplement. 4) Discussion is fine with the current data but complete overhaul suggested above will change the way discussion happens.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1 Comments

R1.1 The Introduction is a bit jumpy – paragraph 1 discusses healthcare resources, poverty, and race, paragraph 2 addresses age, pre-existing conditions, population density, living arrangements, and

poverty again. Organizing along the themes of biologic, demographic, and socioeconomic risk factors, recognizing the overlaps, might improve the structure of the Introduction.
a. Thank you for this suggestion. We have restructured our analysis to focus on individual, household, and community factors instead of biologic, demographic, and socioeconomic factors. We have updated the introduction to reflect this change. We further emphasize the revised structure of the analysis in the Methods section.
b. We have selected one example risk factor within each of these three themes (i.e. age, poverty, group quarters), as well as the cross-cutting factor of proportion of the population that is people of color (R1.3). We have created bivariate maps for each pairwise combination of risk factor categories, for a total of 6 bivariate maps.
c. Lastly, we have also updated citations in the introduction to incorporate the most recent literature.
R1.2 Why the analysis is done at the county level remains unclear. What decisions are made at the county vs. state level, and why is this important? Social distancing policies are mentioned, but those decisions are typically at the state level. The Results are also described at the state level (e.g., Texas, Georgia, Mississippi) While locally tailoring approaches is logical, motivation and interpretation for this particular level of analysis is missing.
a. We thank the reviewer for this suggestion and have clarified our decision to present results on the county-level: “Counties often have flexibility in determining the stringency of their COVID-19 response relative to their respective state orders (38,39); therefore, counties represent a spatial and administrative unit ideal for localized response.”
R1.3 The paper starts and ends with a discussion of race (first paragraph of Introduction and Conclusion), but this focus isn’t clear throughout the paper where race appears in only one of the bivariate analyses. Important intersections of race with several of the other risk factors could have been examined to strengthen this narrative, and would have been very interesting. However, it seems the purpose of this paper is to present examples of intersecting risk, which is also valuable. Clarification on why these five analyses are presented, instead of other potential analyses, would be helpful.
a. We agree that race and ethnicity should be a key part of our analysis. To this end, we have restructured the paper (R1.1) to emphasize the cross-cutting nature of race and ethnicity, and its central importance to discussions of COVID-19, by examining the intersection of race with factors that fall in the other categories – race and ethnicity with the individual factor age; race and ethnicity with the household factor of poverty; and race and ethnicity with the community factor group quarters. We now present three bivariate maps involving race and ethnicity as Figure 2.
b. Our rationale for choosing the 6 bivariate maps now more closely follows the restructured paper (R1.1) and has been clarified in the Methods section: “As an illustrative example, we examine the different pairwise intersections of age (an individual characteristic), poverty (a household characteristic), and prevalence of group homes (a community characteristic) in counties across the US. We also examine how these factors intersect with the proportion of the population that is people of color (i.e., population other than non-Hispanic white), a metric that reflects histories of US race relations.”
Reviewer 1 Minor Comments:
R1.4 Page 5, lines 24-27: Overall, it’s unclear why these are the five analyses presented (out of 24 available risk factors). Adding numbers to the separate analyses could help clarify this sentence.
a. We thank the reviewer for this suggestion. We have reframed the analysis (R1.1) and focused on one example variable from each category (individual, household, community), in addition to the proportion of population that is people of color (R1.3), and created six pairwise comparisons of these four county-level characteristics. We now present these maps in two figures.
R1.5 Page 6, line 12: Why wasn’t the public (i.e. county-level decision makers) involved, especially as the tool is built for them?
a. We thank the reviewer for making this important point. Several authors are members (two are co-founders) of the Covid-19 Mobility Data Network, a collaboration of researchers and public health officials across the US. As such, we have been in nearly-daily communication with public officials in

city and state governments throughout the epidemic. The need for this tool and data stems from our own close interactions with these officials. While none of the county officials meet the criteria for authorship, their input guided the development of this tool.

R1.6 Page 7, lines 2-4: New Mexico, Arizona, and Colorado appear to have the highest concentration of counties with high poverty and populations over 70 (have the most dark purple shading), but aren't mentioned, whereas Georgia, Texas, and Arkansas are highlighted.

a. Thank you for this observation. In the Results section, we have added the proportion of counties in the state that fall in the high-high (dark purple) category to account for the different numbers of counties in different states. We also report the population in the affected counties to provide further context.

R1.7 Page 10, lines 14-18: What's the standard for bed capacity per 100,000?

a. We added summary statistics of county-level bed capacity, "The median county has approximately 185 hospital beds per 100,000 population (mean: 294; IQR: 69, 357; Figure S5)."

b. In addition, we highlight the urban/rural inequality in bed capacity, "According to a May 2020 report from the US Society for Critical Care Medicine, only one percent (963) of all ICU beds are located in rural areas (48)."

c. Lastly, we now add bivariate maps of bed capacity and the covariates of interest in the Supplemental Materials.

R1.8 Page 14, lines 31-35: Useful proposal of modifying these risk factors through financial aid and healthcare. Is there any evidence of financial aid and improved social distancing?

a. This proposal is based on previous research indicating financial incentives to lower "presenteeism" (i.e., going to work while sick) have been shown to be effective with influenza (e.g., 10.2105/AJPH.2013.301269), and the comprehensive stimulus packages of other countries that provide similar financial incentives (e.g., Hong Kong and Japan). However, after restructuring the paper (R1.1), we no longer refer to modifying these risk factors through financial aid.

R1.9 There are a few copy-editing issues, i.e. "U.S." versus "US" versus "United States," missing period on page 12, ect.

a. We thank the reviewer for identifying these issues. We have proofread the manuscript carefully and addressed these copy-editing issues.

Reviewer 2 Comments

R2.1 Always start with references and work which is peer reviewed and published. There is a lot of referencing of data which is from news media which might be biased and unscientific. Speaking of references, please reference introduction, methods and discussion. Results section should talk only about original data.

a. We thank the reviewer for this suggestion. At the time of writing in mid-March, there was a deficit of scientific literature on the topic of social determinants of COVID-19 in the United States, which necessitated using news media articles. We have updated all citations to include the latest peer-reviewed articles on these topics.

b. We added these citations in the Introduction when introducing the county-level characteristics examined in our analysis, as well as in the Discussion section.

R2.2 Methods: I see a lot of maps which can be drawn using CDC public website. There is not much of science here. I would however had been interested in a 2*2 figure where in number of positive cases of COVID would have been put together with these variables/social determinant. Merely looking at current maps are thought provoking but not scientific enough. Along the same lines some amount of basic logistic and linear regression modeling would improve ability to draw conclusions.

a. We chose factors based on existing literature that shows these are important risk factors for COVID-19 susceptibility and severity. Our goal was not to further validate these known risk factors of heightened vulnerability, but to illustrate the inter-county variation that will result in populations being differently affected and will require a highly contextualized response. We believe adding regression results would distract from the main goal of our paper: to provide a tool to evaluate county-specific risks when planning for re-opening and for subsequent waves.

b. We thank the reviewer for this feedback and believe it is an important step for future study.

However, we also believe it is outside the scope of this paper. We now note this explicitly in the conclusion: “Currently, there is insufficient evidence to justify assigning importance weights to different risk factors; however, as more data become available, future research may expand on our analysis by, for example, constructing and evaluating a polysocial risk score (49).”

R2.3 Results: There is too many things going in the results section to draw somewhat homogeneous conclusion. I recommend taking the high impact message and putting it in first and then presenting smaller facts fully in the supplement.
a. Thank you for this suggestion. We have reframed the analysis (R1.1) and streamlined the results to focus on the pairwise intersection of four factors, one from each category (individual, household, community) as well as the proportion of the population that is people of color, as key examples of how these data can be used.
R2.4 Discussion is fine with the current data but complete overhaul suggested above will change the way discussion happens.
a. We have updated the discussion based on the latest available information to focus on how examination of the intersection of these county-level characteristics may be useful when preparing for future waves of COVID-19.