Commentary

Reproducible Science Is Vital for a Stronger Evidence Base During the COVID-19 Pandemic

Karla Therese L. Sy\textsuperscript{1,2}, Laura F. White\textsuperscript{3}, Brooke E. Nichols\textsuperscript{2,4,5}

\textsuperscript{1}Department of Epidemiology, Boston University School of Public Health, Boston, MA USA, \textsuperscript{2}Department of Global Health, Boston University School of Public Health, Boston, MA USA, \textsuperscript{3}Department of Biostatistics, Boston University School of Public Health, Boston, MA USA, \textsuperscript{4}Health Economics and Epidemiology Research Office, Department of Internal Medicine, School of Clinical Medicine, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa, \textsuperscript{5}Department of Medical Microbiology, Amsterdam University Medical Center, Amsterdam, The Netherlands

Reproducible research becomes even more imperative as we build the evidence base on SARS-CoV-2 epidemiology, diagnosis, prevention, and treatment. In his study, Paez assessed the reproducibility of COVID-19 research during the pandemic, using a case study of population density. He found that most articles that assess the relationship of population density and COVID-19 outcomes do not publicly share data and code, except for a few, including our paper, which he stated “illustrates the importance of good reproducibility practices”. Paez recreated our analysis using our code and data from the perspective of spatial analysis, and his new model came to a different conclusion. The disparity between our and Paez’s findings, as well as other existing literature on the topic, give greater impetus to the need for further research. As there has been near exponential growth of COVID-19 research across a wide range of scientific disciplines, reproducible science is a vital component to produce reliable, rigorous, and robust evidence on COVID-19, which will be essential to inform clinical practice and policy in order to effectively eliminate the pandemic.

Importance of reproducible research

Reproducible science, which refers to the ability to achieve identical findings with the same code and data, is essential to critically examine scientific evidence through the replication of results. This in contrast with replicability, which is the act of repeating a study without the original use of data, but generally using the same methods (Broman et al. 2017). The ability to reproduce research results can be used to assess whether different assumptions and models can lead

[Correction added on 16 November 2021, after first online publication: the text ‘redacted’ changed to ‘Sy, White, and Nichols’ in page 2 and 3]

Correspondence: Brooke E. Nichols, Department of Global Health, Boston University School of Public Health, 801 Massachusetts Avenue, Boston, MA 02118, USA
e-mail: brooken@bu.edu

Submitted: 15 October 2021; Revised version accepted: 8 November 2021
Geographical Analysis

to different conclusions, and examine inconsistent results across studies. Reproducible science fosters greater collaboration across disciplines, improves scientific discourse, and strengthens scientific evidence. However, in recent years, the scientific community has established that there is a “reproducibility crisis” (Sayre and Riegelman 2019). In 2016, a survey of researchers found that more than 70% have tried and failed to reproduce another scientist’s experiments, and more than 50% failed to recreate their own (Baker 2016). This lack of scientific rigor and robustness has consequences for the reliability of research findings, and the extent to which the public should trust research.

In particular, during the current COVID-19 pandemic, there has been near exponential growth of COVID-19 literature and evidence across a wide range of scientific disciplines (Brainard 2020). Reproducible research becomes even more imperative as we build the evidence base on SARS-CoV-2 epidemiology, diagnosis, prevention, and treatment. In his study, Paez assessed the reproducibility of COVID-19 research during the pandemic, using a case study of population density (Paez 2021). He found that most publications that assessed the association of population density and COVID-19 outcomes fell short of the gold standard of reproducible research. Currently, the literature for the association of population density and COVID-19 outcomes remains inconclusive, as several studies in various geographic settings have come to opposing conclusions. He found that different researchers used a range of statistical techniques that could account for these varying findings. Reproducibility of studies to verify findings or reanalyze the data could provide greater insight into why associations differ across studies.

Unfortunately, Paez found that most articles that assess the relationship of population density and COVID-19 outcomes do not publicly share data and code, except for a few, including our article (Sy and White 2021), which he stated “illustrates the importance of good reproducibility practices”. He was able to recreate our analysis using our code and data. Additionally, he also reanalyzed our data from the perspective of spatial analysis, and his new model came to a different conclusion.

The perspective that Paez introduces to our work is an example of the importance of reproducibility. He empirically demonstrated that scientific results can have different conclusions depending on model choice and specifications, and that the discrepancy between his and our analysis likely stemmed from different modeling choices, which would not have been apparent had we not made our code and data available. It is well-known that variation in statistical models and assumptions used can vastly change the results, and analytic choices usually depend on the scientists’ field of research. When 29 research groups were asked to answer a research question on the same data set, the effect sizes ranged from OR = 0.89 to OR = 2.83 (Silberzahn et al. 2018). Reproducible science also helps scientists across different fields work together in order to address various knowledge or method gaps, which promotes and fosters more rigorous interdisciplinary research. In this case, his expertise in spatial analysis added to our expertise in infectious disease epidemiology, and offered an alternative conclusion to our findings.

Comparison of Paez’s model and Sy, White, and Nichol model, and other published literature

In his article, Paez noted that the mixed linear models in our analysis were an appropriate modeling choice, but indicated two potential limitations that he attempted to address with spatial models. He addressed the potential non-random sample selection with Heckman’s selection model with spatial filtering, and replaced the log-transformation of population density with a quadratic expansion. In
our study, we conducted several sensitivity analyses to assess the robustness of our results to address any potential limitations in our model; however, we state that sampling could be an issue, “we had to only include counties that had sufficient case data in order to accurately estimate $R_0$”. When Heckman’s selection model was used, Paez found that “the coefficient for population density is still positive, but the magnitude changes: in effect, it appears that the effect of density is more pronounced than what [Sy, White, and Nichols] Model 3 indicated.”, which corresponds to what we predicted in our discussion section “if we included all counties, the true association between population density and $R_0$ would likely be greater than what we report in our analysis given our findings that the counties excluded in the analysis had a significantly lower density and expected very low $R_0$ due to lack of cases.” His selection model confirmed our prediction about the direction of the potential selection bias. Moreover, decisions in how variables are operationalized could also be a reason for different conclusions. Paez’s method of transforming the population density variable allowing for non-monotonic changes to the relationship between population density and the basic reproductive number offered an alternative conclusion, where higher density is not always associated with greater risk of disease spread. The disparity in findings between our and Paez’s findings give greater impetus to the need for further research on this topic.

As noted in our article, “recent literature [on density and COVID-19] has been conflicting, where some research also suggests a density-dependence of COVID-19 transmission (Rocklöv and Sjödin 2020; Rubin et al. 2020) and other measures of the severity of the outbreak (Anand et al. 2020; Wong and Li 2020), while other research suggests that there are other factors that can better explain the pandemic (Hamidi, Ewing, and Sabouri 2020; Hamidi, Sabouri, and Ewing 2020)”. Aside from differences in model specifications, other potential reasons for the variation in conclusions may be due to (a) unmeasured confounding, (b) data quality issues that cause misclassification, (c) selection bias, or (d) other unknown biases. These sources of bias are ubiquitous in observational studies, and researchers do their best to mitigate and limit these potential sources of bias that obscure the true association between exposure and outcome. Reproducible research provides an avenue to critically examine published study results for potential sources of bias that were not properly accounted for.

Conclusion

The COVID-19 pandemic is an unprecedented time for scientific research, with experts from various fields working together to rapidly improve our understanding of SARS-CoV-2. Reproducible science is a vital component to produce reliable, rigorous, and robust literature on COVID-19, which will be essential to inform clinical practice and policy in order to effectively eliminate the pandemic.

References

Anand, S., M. Montez-Rath, J. Han, J. Bozeman, R. Kerschmann, P. Beyer, J. Parsonnet, and G. M. Chertow. (2020). “Prevalence of SARS-CoV-2 Antibodies in a Large Nationwide Sample of Patients on Dialysis in the USA: A Cross-sectional Study.” The Lancet 396(10259), 1335–44. https://doi.org/10.1016/S0140-6736(20)32009-2

Baker, M. (2016). “1,500 Scientists Lift the Lid On Reproducibility.” Nature 533(7604), 452–4. https://doi.org/10.1038/533452a

Brainard, J. (2020). “Scientists are Drowning in COVID-19 Papers. Can New Tools Keep Them Afloat?” Science. Retrieved from https://www.sciencemag.org/content/article/scientists-are-drowning-covid-19-papers-can-new-tools-keep-them-afloat
Broman, K., M. Cetinkaya-Rundel, A. Nussbaum, C. Paciorek, R. Peng, D. Turek, and H. Wickham. (2017). Recommendations to Funding Agencies for Supporting Reproducible Research. Retrieved from https://www.amstat.org/asa/files/pdfs/POL-ReproducibleResearchRecommendations.pdf

Hamidi, S., R. Ewing, and S. Sabouri. (2020). “Longitudinal Analyses of the Relationship Between Development Density and the COVID-19 Morbidity and Mortality Rates: Early Evidence from 1,165 Metropolitan Counties in the United States.” *Health Place* 64, 102378. https://doi.org/10.1016/j.healthplace.2020.102378

Hamidi, S., S. Sabouri, and R. Ewing. (2020). “Does Density Aggravate the COVID-19 Pandemic?” *Journal of the American Planning Association* 86(4), 495–509. https://doi.org/10.1080/01944363.2020.1777891

Paez, A. (2021). “Reproducibility of Research During COVID-19: Examining the Case of Population Density and the Basic Reproductive Rate from the Perspective of Spatial Analysis.” *Geographical Analysis.*

Rocklöv, J., and H. Sjödin. (2020). “High Population Densities Catalyse the Spread of COVID-19.” *Journal of Travel Medicine* 27(3), taaa038. https://doi.org/10.1093/jtm/taaa038

Rubin, D., J. Huang, B. T. Fisher, A. Gasparini, V. Tam, L. Song, X. I. Wang, J. Kaufman, K. Fitzpatrick, A. Jain, H. Griffis, K. Crammer, J. Morris, and G. Tasián. (2020). “Association of Social Distancing, Population Density, and Temperature With the Instantaneous Reproduction Number of SARS-CoV-2 in Counties Across the United States.” *JAMA Network Open* 3(7), e2016099. https://doi.org/10.1001/jamanetworkopen.2020.16099

Sayre, F., and A. Riegelman. (2019). “Replicable Services for Reproducible Research: A Model for Academic Libraries.” *College & Research Libraries* 80(2), 260–272. https://doi.org/10.5860/crl.80.2.260

Silberzahn, R., E. L. Uhlmann, D. P. Martin, P. Anselmi, F. Aust, E. Awtrey, Š. Bahník, F. Bai, C. Bannard, E. Bonnier, R. Carlsson, F. Cheung, G. Christensen, R. Clay, M. A. Craig, A. Dalla Rosa, L. Dam, M. H. Evans, I. Flores Cervantes, N. Fong, M. Gamez-Djokic, A. Glenz, S. Gordon-McKeon, T. J. Heaton, K. Hederos, M. Heene, A. J. Hofelich Mohr, F. Högden, K. Hui, M. Johannesson, J. Kalodimos, E. Kaszubowski, D. M. Kennedy, R. Lei, T. A. Lindsay, S. Liverani, C. R. Madan, D. Molden, E. Molleman, R. D. Morey, L. B. Mulder, B. R. Nijstad, N. G. Pope, B. Pope, J. M. Prenoveau, F. Rink, E. Robusto, H. Roderique, A. Sandberg, E. Schlüter, F. D. Schönbrodt, M. F. Sherman, S. A. Sommer, K. Sotak, S. Spain, C. Spörlein, T. Stafford, L. Stefanutti, S. Tauber, J. Ullrich, M. Vianello, E.-J. Wagenmakers, M. Witkowiak, S. Yoon, and B. A. Nosek. (2018). “Many Analysts, One Data Set: Making Transparent How Variations in Analytic Choices Affect Results.” *Advances in Methods and Practices in Psychological Science* 1(3), 337–56. https://doi.org/10.1177/2515245917747646

Sy, K. T. L., L. F. White, and B. E. Nichols. (2021). “Population Density and Basic Reproductive Number of COVID-19 Across United States Counties.” *PLoS One* 16(4), e0249271. https://doi.org/10.1371/journal.pone.0249271

Wong, D. W. S., and Y. Li. (2020). “Spreading of COVID-19: Density Matters.” *PLoS One* 15(12), e0242398. https://doi.org/10.1371/journal.pone.0242398