

localcovid19now: processing and mapping COVID-19 case data at subnational scales

Stephen J. Beckett ¹✉, Freyja A. Brandel-Tanis ^{2,3}, Quan Nguyen^{1,4},
Aroon T. Chande ⁵, Lavanya Rishishwar ^{1,5}, Clio Andris ^{2,6}, and Joshua
S. Weitz ^{1,7,8}

1 School of Biological Sciences, Georgia Institute of Technology, Atlanta, GA, United States of America
2 School of City and Regional Planning, Georgia Institute of Technology, Atlanta, GA, United States of America
3 School of Civil and Environmental Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America
4 School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA, United States of America
5 Applied Bioinformatics Laboratory, Georgia Institute of Technology, Atlanta, GA, United States of America
6 School of Interactive Computing, Georgia Institute of Technology, Atlanta, GA, United States of America
7 School of Physics, Georgia Institute of Technology, Atlanta, GA, United States of America
8 Institut de Biologie, École Normale Supérieure, Paris, France ✉ Corresponding author

DOI: [10.21105/joss.04898](https://doi.org/10.21105/joss.04898)

Software

- [Review](#) 
- [Repository](#) 
- [Archive](#) 

Editor: [Gabriela Alessio Robles](#) 



Reviewers:

- [@richelbilderbeek](#)
- [@mbkumar](#)
- [@welch16](#)
- [@Epic19mz](#)

Submitted: 22 September 2022

Published: 21 January 2023

License

Authors of papers retain copyright and release the work under a Creative Commons Attribution 4.0 International License ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).

Summary

The [localcovid19now](#) R package provides functionality to load, unify and visualize recent COVID-19 case data at subnational scales in order to provide localized situational reports and improve understanding of the scale of local COVID-19 transmission. The package loads data from a variety of data sources and returns the most recent estimate of recorded per capita active COVID-19 infections, the date of the most recent report, and the geometry of each region. These data can then be visualized via mapping documented per capita active infections. We also provide functionality to visualize the risk of exposure to COVID-19, given a particular event size.

Statement of need

The COVID-19 global pandemic has led to hundreds of millions of recorded cases, over six million recorded deaths, and critical pressures on health and economic systems ([World Health Organization, 2022](#)). A key challenge in mitigating ongoing spread is that disease transmission and human behavior are coupled ([Bavel et al., 2020](#); [Funk et al., 2010](#); [Weitz et al., 2020](#)). Coupled viral and human dynamics have limited the predictive window of infectious disease models (including those of COVID-19), while also catalyzing efforts to develop awareness campaigns that leverage localized real-time information to inform the public and mitigate transmission. Beginning in May 2020, our team launched the COVID-19 Event Risk Tool (<https://covid19risk.biosci.gatech.edu/>), a situational dashboard with maps estimating exposure risk to COVID-19 across counties in the US ([Chande et al., 2020](#)). We use the number of active cases per capita and an assumed case ascertainment bias to estimate current local COVID-19 prevalence p . We then quantify and visualize COVID-19 exposure risk κ (%) within a group of size n as $\kappa = 100 \times (1 - (1 - p)^n)$. Since launch, more than 16M individuals have accessed the site and several similar international efforts have been initiated, e.g., for [Italy](#) ([Eventi e Covid-19, 2020](#)), [Spain](#) ([Eventos y Covid-19, 2020](#)), [Mexico](#) ([COVID-19 México, 2020](#)), [Argentina](#) ([Cassese, 2020](#)). Here, in the *localcovid19now* R package, we extend our initial release ([Chande et al., 2020](#)) to develop a pipeline to combine population estimates, COVID-19 case data, and spatial geometry such that risk estimates can

be performed at localized scales across multiple countries. Our package aims to facilitate the process of downloading and visualizing recent COVID-19 data, in support of public health research and campaigns to inform and alert the public of the ongoing risk of transmission – and is of particular interest to public health scientists and researchers, epidemiologists, community leaders, and policymakers.

Our work is complementary to existing efforts to provide localized COVID-19 data across multiple countries, e.g., ([Dong et al., 2020](#); [Guidotti & Ardia, 2020](#); [Palmer et al., 2021](#); [Wahlteiz & others, 2020](#)). Via *localcovid19now*, we focus on aggregating recent case data rather than assembling a timeseries of all different COVID-19 data metrics. We use recent case data as the basis for providing localized real-time estimates of active cases. Additionally, spatial coverage differs among existing aggregate efforts to harmonize COVID-19 data. In order to maximize geographic coverage, the *localcovid19now* package pulls in data collated by John Hopkins U CSSE ([Dong et al., 2020](#)), from Google COVID-19 Open Data ([Wahlteiz & others, 2020](#)), the Humanitarian Data Exchange, the COVID-19 Data Hub Guidotti ([2022](#)), and the [WHO European Region COVID19 Subnational Explorer](#) ([WHO-EURO COVID19 Subnational Explorer, 2020](#)). We also include additional governmental/volunteer curated subnational datasets, e.g., ([Berry et al., 2021](#); [Cota, 2020](#); [Jefferies et al., 2020](#)), (a full list of sources is provided with the [package](#)). Individual sources store data in different file formats, languages and file structures requiring tailored data retrieval strategies. Given that as time progresses and needs or efforts of those maintaining COVID-19 datasets change, the sourced datasets may change formats (e.g., Czechia added an extra identifier column to their dataset in November 2021 ([Komenda et al., 2020](#))), change reporting areas (e.g., administrative border changes in the Netherlands), become deprecated (e.g., Cuba which last updated in September 2021), or otherwise be forced offline (e.g., [Ireland](#) ([McNamee, 2021](#))). For this reason we have modularized our data reading process such that errors in one part of the pipeline do not prevent the loading of other datasets. An example of the number of active cases per 100,000 individuals as read and visualized via the *localcovid19now* pipeline is shown in [Figure 1](#).

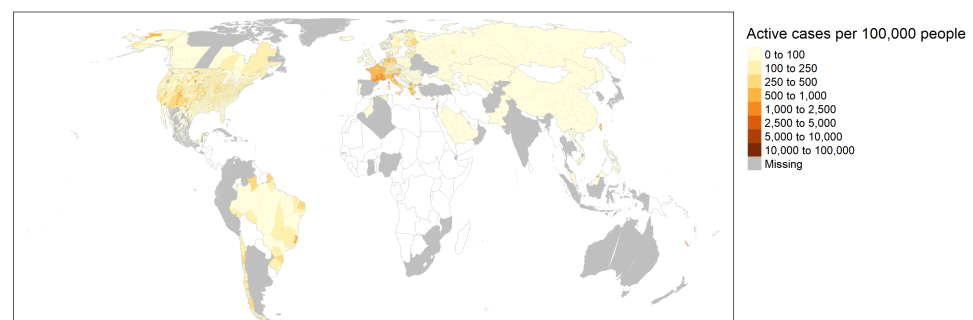


Figure 1: Active cases per 100,000 people (as of 9 December 2022). Greyscale polygons denote areas where data estimates are over 30 days old, which we show here to show the current package scope. This map was created using the *localcovid19now* package by running the code: `GlobalMap = LoadData(dropNACountries=FALSE); PerCapitaMap_tmap(GlobalMap,100000)`. More examples are shown on the packages homepage.

Our motivation in developing this project is to provide data-driven localized epidemic information that connects case data to potential exposure risk. While individual perception of COVID-19 exposure risk is linked to the utilization of risk-reducing actions, individual assessment of risk has been found to be generally uncorrelated to estimated exposure risk ([Sinclair, Hakimi, et al., 2021](#)). In contrast, communicating infection exposure risk estimates, as we do at the [COVID-19](#)

[Event Risk Tool](#), has been shown as an effective intervention in realigning perceptions of localized event risks ([Sinclair, Hakimi, et al., 2021](#); [Sinclair, Stanley, et al., 2021](#); [Sinclair et al., 2022](#)). The package additionally includes functionality to help users visualize data (via leaflet ([Cheng et al., 2021](#)) or tmap ([Tennekes, 2018](#))), including that of per capita active documented infections and estimated event-based risk.

Acknowledgements

We thank Seolha Lee for her contributions to developing the COVID-19 Event Risk Tool. We thank Kaitlyn Johnson, Robel Kassa, and Leo Wolansky for constructive feedback. Development of the [localcovid19now](#) package was supported by funds from the Centers for Disease Control and Prevention (75D30121P10600), The Rockefeller Foundation (11356), the Charities in Aid Foundation, the Marier Cunningham Foundation, and the Chaire Blaise Pascal funds of the Île-de-France region.

Author contributions

SJB led the project. SJB, FAB, QN and ATC contributed to dataset collection and coding methods for data curation. SJB, FAB and ATC contributed to package development. LR, CA, and JSW provided substantial contributions via decision and design guidance, and project management. SJB wrote the manuscript with contributions from all authors.

References

- Bavel, J. J. V., Baicker, K., Boggio, P. S., Capraro, V., Cichocka, A., Cikara, M., Crockett, M. J., Crum, A. J., Douglas, K. M., Druckman, J. N., & others. (2020). Using social and behavioural science to support COVID-19 pandemic response. *Nature Human Behaviour*, 4(5), 460–471. <https://doi.org/10.1038/s41562-020-0884-z>
- Berry, I., O'Neill, M., Sturrock, S. L., Wright, J. E., Acharya, K., Brankston, G., Harish, V., Kornas, K., Maani, N., Naganathan, T., & others. (2021). A sub-national real-time epidemiological and vaccination database for the COVID-19 pandemic in Canada. *Scientific Data*, 8(1), 1–10. <https://doi.org/10.1038/s41597-021-00955-2>
- Cassese, N. (2020). *Fiestas de fin de año. Calculá el riesgo de cruzarte un infectado según tu lugar de festejo y la cantidad de invitados.* (accessed September 8, 2022) <https://www.lanacion.com.ar/sociedad/fiestas-fin-ano-calcula-riesgo-coronavirus-infectado-nid2548498/#/>.
- Chande, A., Lee, S., Harris, M., Nguyen, Q., Beckett, S. J., Hilley, T., Andris, C., & Weitz, J. S. (2020). Real-time, interactive website for US-county-level COVID-19 event risk assessment. *Nature Human Behaviour*, 4(12), 1313–1319. <https://doi.org/10.1038/s41562-020-01000-9>
- Cheng, J., Karambelkar, B., & Xie, Y. (2021). *Leaflet: Create interactive web maps with the JavaScript 'leaflet' library.* <https://CRAN.R-project.org/package=leaflet>
- Cota, W. (2020). Monitoring the number of COVID-19 cases and deaths in Brazil at municipal and federative units level. *SciELO Preprints*. <https://doi.org/10.1590/SciELOPreprints.362>
- COVID-19 México. (2020). (accessed September 8, 2022) https://adrian-acuna.shinyapps.io/Dashboard-MX-v3/_w_33ffcba1/#shiny-tab-Riesgo.
- Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases*, 20(5), 533–534. [https://doi.org/10.1016/S1473-3099\(20\)30120-1](https://doi.org/10.1016/S1473-3099(20)30120-1)

- Eventi e covid-19*. (2020). (accessed September 8, 2022) <https://covid19eventi.datainterfaces.org/>.
- Eventos y covid-19*. (2020). (accessed September 8, 2022) <https://eventosycovid19.es/>.
- Funk, S., Salathé, M., & Jansen, V. A. (2010). Modelling the influence of human behaviour on the spread of infectious diseases: A review. *Journal of the Royal Society Interface*, 7(50), 1247–1256. <https://doi.org/10.1098/rsif.2010.0142>
- Guidotti, E. (2022). A worldwide epidemiological database for COVID-19 at fine-grained spatial resolution. *Scientific Data*, 9(1), 112. <https://doi.org/10.1038/s41597-022-01245-1>
- Guidotti, E., & Ardia, D. (2020). COVID-19 Data Hub. *Journal of Open Source Software*, 5(51), 2376. <https://doi.org/10.21105/joss.02376>
- Jefferies, S., French, N., Gilkison, C., Graham, G., Hope, V., Marshall, J., McElnay, C., McNeill, A., Muellner, P., Paine, S., & others. (2020). COVID-19 in New Zealand and the impact of the national response: a descriptive epidemiological study. *The Lancet Public Health*, 5(11), e612–e623. [https://doi.org/10.1016/S2468-2667\(20\)30225-5](https://doi.org/10.1016/S2468-2667(20)30225-5)
- Komenda, M., Panoška, P., Bulhart, V., Žofka, J., Brauner, T., Hak, J., Jarkovský, J., Mužík, J., Blaha, M., Kubát, J., Klimeš, D., Langhammer, P., Daňková, Š., Májek, O., Bartůňková, M., & Dušek, L. (2020). *COVID-19: Přehled aktuální situace v ČR. Onemocnění aktuálně* [online]. Praha: Ministerstvo zdravotnictví ČR, 2020 [cit. 09.05.2022]. <https://onemocneni-aktualne.mzcr.cz/covid-19>
- McNamee, M. S. (2021). *CHSE cyber-attack: Irish health service still recovering months after hack*. (accessed September 8, 2022) <https://www.bbc.com/news/world-europe-58413448>.
- Palmer, J., Sherratt, K., Martin-Nielsen, R., Bevan, J., Gibbs, H., Funk, S., Abbott, S., Group, C. C. W., & others. (2021). Covidregionaldata: Subnational data for COVID-19 epidemiology. *Journal of Open Source Software*, 6(63), 3290. <https://doi.org/10.21105/joss.03290>
- Sinclair, A. H., Hakimi, S., Stanley, M. L., Adcock, R. A., & Samanez-Larkin, G. R. (2021). Pairing facts with imagined consequences improves pandemic-related risk perception. *Proceedings of the National Academy of Sciences*, 118(32). <https://doi.org/10.1073/pnas.2100970118>
- Sinclair, A. H., Stanley, M. L., Hakimi, S., Cabeza, R., Adcock, R. A., & Samanez-Larkin, G. R. (2021). Imagining a personalized scenario selectively increases perceived risk of viral transmission for older adults. *Nature Aging*, 1(8), 677–683. <https://doi.org/10.1038/s43587-021-00095-7>
- Sinclair, A. H., Taylor, M., Brandel-Tanis, F., Davidson, A., Chande, A. T., Rishishwar, L., Andris, C. M., Adcock, R. A., Weitz, J. S., Samanez-Larkin, G., & Beckett, S. J. (2022). Real-time interventions counteract COVID-19 risk misestimation in the United States. *PsyArXiv*. <https://doi.org/10.31234/osf.io/v8tdf>
- Tennekes, M. (2018). tmap: Thematic maps in R. *Journal of Statistical Software*, 84(6), 1–39. <https://doi.org/10.18637/jss.v084.i06>
- Wahlteiz, O., & others. (2020). *COVID-19 Open-Data: curating a fine-grained, global-scale data repository for SARS-CoV-2*. <https://goo.gl/covid-19-open-data>
- Weitz, J. S., Park, S. W., Eksin, C., & Dushoff, J. (2020). Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *Proceedings of the National Academy of Sciences*, 117(51), 32764–32771. <https://doi.org/10.1073/pnas.2009911117>
- WHO-EURO COVID19 subnational explorer. (2020). (accessed September 8, 2022) <https://experience.arcgis.com/experience/3a056fc8839d47969ef59949e9984a71>.

World Health Organization. (2022). *COVID-19 weekly epidemiological update, edition 107, 31 August 2022* (pp. 11 p.) [Technical documents]. World Health Organization.