Article Summary Line: We identified, visualized, and estimated the speed and spatial extent of

the two largest US SARS-CoV-2 waves over 2020-2021.

Running Title: Dynamics of SARS-CoV-2 infection waves in the US.

Keywords: Epidemiology, spatial epidemiology, infectious diseases, emerging infectious

disease, infectious disease spreading.

Title: Quantifying the spatiotemporal dynamics of the first two epidemic waves of SARS-CoV-2

infections in the United States

Authors: Rafael Lopes[1,2,*,#], Nicole A. Swartwood[3,*,#], Yu Lan[1,2], Melanie H. Chitwood[1,2],

Fayette Klaassen[3], Joshua A. Salomon[4], Nicolas A. Menzies[3], Joshua L. Warren[5,2], Nathan D.

Grubaugh[1,2], & Ted Cohen[1,2,†]

Affiliations:

[1] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, Yale

University, New Haven, Connecticut, USA (R. Lopes, Y. Lan, M. Chitwood, N. Grubaugh, T.

Cohen)

[2] Public Health Modeling Unit, Yale School of Public Health, Yale University, New Haven,

Connecticut, USA (R. Lopes, Y. Lan, M. Chitwood, N. Grubaugh, T. Cohen)

[3] Department of Global Health and Population, Harvard T.H. Chan School of Public Health,

Harvard University, Boston, Massachusetts, USA (N. Swartwood, F. Klaassen, N. Menzies)

[4] Department of Medicine, Stanford University, Stanford, California, USA (J. Salomon)

[5] Department of Biostatistics, Yale School of Public Health, Yale University, New Haven,

Connecticut, USA (J. Warren)

* Co-first authors

† Senior author

Corresponding Authors:

Rafael Lopes, 350 George Street, New Haven, CT 06511.

Email: rafael.lopes@yale.edu Phone (+1) 475-300-8845

Nicole A. Swartwood, 90 Smith Street, Boston, MA 02120.

Email: nswartwood@hsph.harvard.edu. Phone (+1) 617-432-6171.

31   Abstract - 122 words

SARS-CoV-2 infection rates displayed striking temporal and spatial variation during the

emergence of new variants globally and within the United States. While spatiotemporal "waves"

of infection have been observed, quantitative assessments of their spread remain limited. Here,

we estimate and compare the speed and spatial extent of the first two major infection waves in

the United States, illustrating these dynamics through detailed visualizations. Our findings reveal

that the origins of these waves coincide with large gatherings and the relaxation of masking

mandates. Notably, the second wave spread more rapidly than the first, driven by multiple, non-

contiguous origins of infection. This highlights the role of regional heterogeneity in epidemic

dynamics and underscores the importance of localized public health measures in mitigating

ongoing outbreaks.

Text – 1900 words

Introduction

Person-to-person transmission of SARS-CoV-2 was confirmed in the United States (US)

in late January 2020 and a national public health emergency was declared six weeks later (1,2).

Spatial visualizations of observed and estimated COVID-19 cases and deaths displayed wave-

like spread as new variants were introduced in the country (3,4). Previous investigations have

examined the correlation between the spatiotemporal patterns of COVID-19 spread and human

mobility (5–7). Other studies have estimated the effectiveness of public health measures such as

lockdowns and travel restrictions in interrupting or modifying these patterns (7–9). However,

efforts to quantify the speed and spatial extent of viral spread over distinct epidemic waves have

been limited. Here, we apply a version of the Besag, York, and Mollié (BYM) spatial model

(10,11) to temporally- and spatially-resolved estimates of SARS-CoV-2 infections derived from

a previously described model (3). Based on this analysis we quantify the speed at which SARS-

CoV-2 outbreaks spread during the two large waves of SARS-CoV-2 infections in the United

States during March 2020 to December 2021.

Methods

Data sources

We used previously published daily estimates of SARS-CoV-2 infections at the county-

level in the United States from a Bayesian nowcasting model that synthesized reported COVID-

19 cases and deaths, accounting for both under ascertainment and time lags (3). Our dataset

encompasses the period from March 2020 to December 2021. To estimate population

denominators for per capita analyses, we used 2019 population estimates from the US Census

Bureau, which provided population size estimates for all Census Block Groups (CBGs) in the

United States (12).

Regularization of geographical units

The smallest consistently available unit of reporting for US COVID-19 cases,

hospitalizations, and deaths was at the county-level, which dictated the geographic resolution of

the SARS-CoV-2 infection estimates. However, as the geographic size and population density of

counties differs systematically across the US (larger and less dense in the West than in the East),

a county-level analysis could lead to bias in estimates of the wave speed and expansion.

To improve our ability to detect spatial patterns in the data, we first created a spatially

regularized grid of infections per capita across the contiguous US. The grid is comprised of

7,665 unique hexagons, each enclosing 64.75 square kilometers (25 square miles). We employed

an area-weighted approach to distribute population estimates from CBGs to hexagons (**Figure**

1A-C), assuming a constant population throughout the analysis period. We then distributed

estimated infections from counties to hexagons (12) based on the fraction of each county's

population contained in each hexagon and assuming that per capita infection rates were

distributed equally within a county. We combined these to produce daily estimates of SARS-

CoV-2 infections per capita for each hexagon (**Figure 1D-E**).

Assessment of infections per capita surfaces

89   We fit a modified version of the BYM model as implemented in R-INLA (10), called

90   BYM2, to estimate spatially smoothed rates of SARS-CoV-2 infections per capita across the

91   hexagonal grid. The BYM2 is a reparametrized version of the original Besag (11) model, which

92   improves the assignment of prior distributions for the corresponding model parameters and the

93   subsequent interpretation of parameters (10). Specifically, our model is given as

$$Z(A_i) = \mu + \theta(A_i) + \epsilon(A_i), \qquad (1)$$

where $Z(A_i)$ is the infections per capita in hexagon A_i, μ is the global intercept, $\theta(A_i)$ are random

effects that are assigned the BYM2 prior distribution, and $\epsilon(A_i)$ are residual error terms assumed

to be statistically independent. For more details on the model and on the priors used see the

Supplementary Material. We fit this model separately for each day over March 2020 to

December 2021. The collection of all the $\mu + \theta(A_i)$ (i.e., the global intercept plus the denoised

and spatially smoothed random effects) are the modeled infections per capita, or surfaces. We

then produced a daily sequence of these surfaces, which we used to calculate the speed of wave

expansion and to visualize changes over time. We opted for a separate model fit on each day to

avoid temporal over smoothing effects, as the infection estimates were derived from a

nowcasting model that already applied temporal smoothing.

Definitions of wave and speed of expansion

We defined waves as one or more sets of contiguous hexagons with relatively high

infections-per-capita levels. To allow for multiple infection centers with the same wave, we did

not limit our definition to a single contiguous set. Using this framework, we categorized 'Wave

1' as the period from September 2020 to February 2021 and 'Wave 2' as the period from July

2021 to November 2021. To formally determine hexagons within a wave, we defined a level of

per-capita SARS-CoV-2 infection, above which we coded hexagons as being part of the infection

wave. For our main analysis we used a threshold of 165 daily infections per-capita to

characterize a hexagon as being within a wave. This value represents the 75th percentile of the

infections-per-capita value distribution during the study period (**Figure S1**). We conducted

sensitivity analyses using 85 and 300 infections per capita as alternative threshold values

(**Supplementary Material**). We calculated the speed of wave expansion as the increase in area

covered by the wave each day (number of hexes recruited into a wave each day).

Statistical analysis

All analyses were performed using R Statistical Software (v4.3.0) (13)). We also used the

INLA, **sf**, **sp**, **spdep**, **areal**, **rgeoda**, and **magick** R packages (14–18).

Results

Spatiotemporal patterns of SARS-CoV-2 infections across the United States

As described in the **Methods**, we categorized infections as belonging to Wave 1

(September 2020 to February 2021) or Wave 2 (July 2021 to December 2021) (**Figure 2A**).

Figure 2B-I shows the smoothed estimates of SARS-CoV-2 infections per capita from the

BYM2 model on eight dates, leading up to the peak of each of the two waves analyzed. Our

model estimated that Wave 1 originated with a set of hexes spanning central South Dakota,

eastern North Dakota, and northeastern Montana (**Figure 2B**) in mid-September of 2020. Then

the wave extended south and to both coasts from this point of origin, achieving a peak of over

1.5 million infections/day by mid-November 2020 (**Figure 2C-E**). Wave 2 originated in the

Ozarks (southern Missouri and northern Arkansas) (**Figure 2F**) in early July 2021, and then

expanded further south (**Figure 2G-I**). Secondary centers of infection appeared later in July in

the Pacific Northwest (**Figure 2G**), and the wave of infections subsequently spread throughout

the western United States (**Figure 2G-I**). This wave had a peak of 300 per capita infection/day

by early September 2021. Animations of the infection waves are provided in **Movie 1,** which

presents maps of the spatially smoothed infections across the full study period at a weekly

timestep.

Wave expansion and speed of expansion

Figure 3B-C shows contour plots of the speed of expansion for each wave. At its peak,

each wave involved the total contiguous United States (8,595,756 km^2), but the two waves

differed in the time needed to spread across the nation. Wave 1 expanded from an area of

327,780 km^2 to the total United States in 64 days (September 8th–November 11th, 2020) (**Figure**

3B). Wave 2 expanded from an area of 246,047 km^2 to the total United States in 51 days (July 7th

– August 27th, 2021) (**Figure 3C**).

Wave 1 reached its maximal speed of expansion (~ 580,000 km^2/day) 49 days after its

formation and 14 days prior to its peak. Wave 2 reached its maximal speed of expansion (~

650,000 km^2/day) more quickly, in only 38 days after its formation and 25 days prior to its peak

(**Figure 3C**). Sensitivity analyses using different thresholds for categorizing a hexagon as within

a wave resulted in consistent estimates for speed of wave expansion and time to reach peak

expansion speed (**Figure S2**). The shape and speed of wave patterns were also robust to the

choice of threshold (**Figure S2** and **S3**).

Discussion

We sought to quantify the expansion of SARS-CoV-2 infections across the United States

during the first two large waves of the epidemic (**Figures 2**, **Movie1**, and **Movie2**). The two

waves had unique origin sites within the United States. We estimated that Wave 1 originated in

parts of northeastern Montana, eastern North Dakota, and central South Dakota in September

2020 (**Figures 2B-E** and **Movie 2**). This estimated origin coincided with a motorcycle rally that

brought over 460,000 individuals to Sturgis, South Dakota in August 2020 and was later

epidemiologically implicated in elevated local COVID-19 rates and interstate spread of infection

(19,20). Wave 2 originated in the Ozarks (**Figures 2F-I** and **Movie 2**), during a time when there

was local re-opening of indoor music venues and relaxation of local masking policies (21,22).

Shortly following spread in the Ozarks, secondary centers of elevated SARS-CoV-2 infections

per capita were found in the Pacific Northwest. Previous work has implicated both domestic and

international travel in contributing to SARS-CoV-2 transmission across United States and likely

explains the rapid appearance of multiple wave centers for the second examined wave of the

epidemic (23).

While we found that Wave 2 had both a higher maximal speed of expansion and achieved

this peak speed more quickly than Wave 1, the overall similarity in our estimates of the speeds of

wave expansion for these first two waves is notable. This similarity in wave speed expansion is

more remarkable in light of the very different viral variants in circulation (Wave 1 was initially

due to wildtype SARS-CoV-2 and later the Alpha variant (24,25), while Wave 2 was due to the

spread of the Delta variant (23)), regional and secular changes in nonpharmaceutical intervention

policies, and the stark differences in the levels and types of infection- and vaccine-induced

population immunity at the time of these two waves (26). Such stability in the speed of wave

expansion suggests that underlying patterns of human mobility may have had a stronger impact

on the patterns of local epidemic expansion than viral variation, non-pharmaceutical

interventions, and host-immunity (7,20).

Our effort to provide quantitative estimates of speeds and locations of epidemic wave

expansion required several simplifying assumptions. Importantly, we elected to use estimates of

infections from a nowcasting model (rather than case notification data) as diagnosis and case

reporting varied markedly in quality and completeness over the epidemic (27). To overcome

potential bias associated with the irregular shapes and sizes of US counties (the smallest unit at

which estimates were possible), we distributed estimated infections on a hexagonal grid before

calculating the speed of wave expansion. This could have introduced bias if cases within

counties were not randomly distributed, though this effect would be modest. Finally, we

arbitrarily selected the threshold of infections per capita to define wave membership. However,

our sensitivity analyses found that the wave speeds were similar at alternative thresholds.

Factors that affect the magnitude and speed of epidemic waves is an area of substantial

interest for other viral diseases like influenza (e.g. Viboud et al. 2006 (28), Eggo et al. 2011 (29),

Gog et al. 2014 (30)). While our work on SARS-CoV-2 does not itself provide explanations for

why each of these first two epidemic waves spread in the manner that they did once they were

established, the development of methods to quantify the speed and extent of spread, such as

those we employ here, are a necessary first step. Further research to explore the manner in which

the pathogen, environment, and host characteristics affect the speeds and patterns of epidemic

expansion may help to predict wave expansion and could provide valuable information for the

planning of spatially defined interventions.

Author Bio (first author only, unless there are only 2 authors)

Rafael Lopes is a mathematical and statistical modeler with interest in methods to predict

outbreak dynamics and infectious disease spread, he holds a Ph.D. in theoretical Physics and

currently is Postdoctoral Associate Researcher at Yale University, Yale School of Public Health,

Department of Epidemiology of Microbial Diseases.

References

1. List of Public Health Emergency Declarations [Internet]. [cited 2024 Dec 3]. Available from:
https://aspr.hhs.gov:443/legal/PHE/Pages/default.aspx

2. Determination that a Public Health Emergency Exists [Internet]. [cited 2024 Dec 3].
Available from: https://aspr.hhs.gov:443/legal/PHE/Pages/2019-nCoV.aspx

3. Chitwood MH, Russi M, Gunasekera K, Havumaki J, Klaassen F, Pitzer VE, et al.
Reconstructing the course of the COVID-19 epidemic over 2020 for US states and counties:
Results of a Bayesian evidence synthesis model. PLOS Comput Biol [Internet]. 2022 Aug 30
[cited 2023 Oct 6];18(8):e1010465. Available from:
https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010465

4. Gamio L, Smith M, Yourish K, Almukhtar S. Watch How the Coronavirus Spread Across
the United States. The New York Times [Internet]. 2020 Mar 21 [cited 2023 Dec 19];
Available from: https://www.nytimes.com/interactive/2020/03/21/us/coronavirus-us-cases-
spread.html

5. Kissler SM, Kishore N, Prabhu M, Goffman D, Beilin Y, Landau R, et al. Reductions in
commuting mobility correlate with geographic differences in SARS-CoV-2 prevalence in
New York City. Nat Commun [Internet]. 2020 Sep 16 [cited 2024 Oct 17];11(1):4674.
Available from: https://www.nature.com/articles/s41467-020-18271-5

6. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and
dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures.
Proc Natl Acad Sci [Internet]. 2020 May 12 [cited 2024 Oct 14];117(19):10484–91.
Available from: https://www.pnas.org/doi/10.1073/pnas.2004978117

7. Kraemer MUG, Yang CH, Gutierrez B, Wu CH, Klein B, Pigott DM, et al. The effect of
human mobility and control measures on the COVID-19 epidemic in China. Science
[Internet]. 2020 May [cited 2024 Oct 2];368(6490):493–7. Available from:
https://www.science.org/doi/full/10.1126/science.abb4218

8. Pepe E, Bajardi P, Gauvin L, Privitera F, Lake B, Cattuto C, et al. COVID-19 outbreak
response, a dataset to assess mobility changes in Italy following national lockdown. Sci Data
[Internet]. 2020 Jul 8 [cited 2024 Oct 2];7(1):230. Available from:
https://www.nature.com/articles/s41597-020-00575-2

9. Chinazzi M, Davis JT, Ajelli M, Gioannini C, Litvinova M, Merler S, et al. The effect of
travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak.
Science [Internet]. 2020 Apr 24 [cited 2024 Oct 2];368(6489):395–400. Available from:
https://www.science.org/doi/full/10.1126/science.aba9757

10. Riebler A, Sørbye SH, Simpson D, Rue H. An intuitive Bayesian spatial model for disease
mapping that accounts for scaling. Stat Methods Med Res [Internet]. 2016 Aug 1 [cited 2024
Oct 2];25(4):1145–65. Available from: https://doi.org/10.1177/0962280216660421

11. Besag J, York J, Molli A. Bayesian image restoration, with two applications in spatial statistics. Ann Inst Stat Math [Internet]. 1991 Mar [cited 2024 Oct 2];43(1):1–20. Available from: http://link.springer.com/10.1007/BF00116466

12. topojson/us-atlas [Internet]. TopoJSON; 2024 [cited 2024 Aug 7]. Available from: https://github.com/topojson/us-atlas

13. R: The R Project for Statistical Computing [Internet]. [cited 2024 Dec 3]. Available from: https://www.r-project.org/

14. Prener C, Revord C. areal: An R package for areal weighted interpolation [Internet]. Zenodo; 2020 [cited 2024 Aug 7]. Available from: https://zenodo.org/records/3822534

15. Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. R J [Internet]. 2018 [cited 2024 Oct 5];10(1):439. Available from: https://journal.r-project.org/archive/2018/RJ-2018-009/index.html

16. magick: Advanced Graphics and Image-Processing in R [Internet]. [cited 2024 Oct 5]. Available from: https://ropensci.r-universe.dev/magick

17. Bivand R. spdep: Spatial Dependence: Weighting Schemes, Statistics [Internet]. 2002 [cited 2024 Oct 7]. p. 1.3-6. Available from: https://CRAN.R-project.org/package=spdep

18. Lindgren F, Rue H. Bayesian Spatial Modelling with R-INLA. J Stat Softw [Internet]. 2015 Feb 16 [cited 2024 Oct 14];63:1–25. Available from: https://doi.org/10.18637/jss.v063.i19

19. Firestone MJ. COVID-19 Outbreak Associated with a 10-Day Motorcycle Rally in a Neighboring State — Minnesota, August–September 2020. MMWR Morb Mortal Wkly Rep [Internet]. 2020 [cited 2024 May 22];69. Available from: https://www.cdc.gov/mmwr/volumes/69/wr/mm6947e1.htm

20. Dave D, McNichols D, Sabia JJ. The contagion externality of a superspreading event: The Sturgis Motorcycle Rally and COVID-19. South Econ J. 2021 Jan;87(3):769–807.

21. Barone E. TIME. 2021 [cited 2024 May 28]. How the Delta Variant Overtook Missouri: A Lesson for the Rest of the U.S. Available from: https://time.com/6085454/delta-variant/

22. Daly M. How a Tiny Town With an Anti-Mask Mayor Caused COVID Chaos. The Daily Beast [Internet]. 2021 Jul 23 [cited 2024 Oct 14]; Available from: https://www.thedailybeast.com/missouris-covid-surge-started-in-tourist-town-of-branson

23. Bolze A, Luo S, White S, Cirulli ET, Wyman D, Dei Rossi A, et al. SARS-CoV-2 variant Delta rapidly displaced variant Alpha in the United States and led to higher viral loads. Cell Rep Med [Internet]. 2022 Mar 15 [cited 2024 Mar 26];3(3):100564. Available from: https://www.sciencedirect.com/science/article/pii/S2666379122000714

24. Alpert T, Brito AF, Lasek-Nesselquist E, Rothman J, Valesano AL, MacKay MJ, et al. Early introductions and transmission of SARS-CoV-2 variant B.1.1.7 in the United States. Cell

[Internet]. 2021 May [cited 2023 May 11];184(10):2595-2604.e13. Available from: https://linkinghub.elsevier.com/retrieve/pii/S0092867421004347

25. Washington NL, Gangavarapu K, Zeller M, Bolze A, Cirulli ET, Barrett KMS, et al. Emergence and rapid transmission of SARS-CoV-2 B.1.1.7 in the United States. Cell [Internet]. 2021 May 13 [cited 2024 Nov 7];184(10):2587-2594.e7. Available from: https://www.cell.com/cell/abstract/S0092-8674(21)00383-4

26. Klaassen F, Chitwood MH, Cohen T, Pitzer VE, Russi M, Swartwood NA, et al. Changes in Population Immunity Against Infection and Severe Disease From Severe Acute Respiratory Syndrome Coronavirus 2 Omicron Variants in the United States Between December 2021 and November 2022. Clin Infect Dis [Internet]. 2023 Aug 1 [cited 2023 Oct 16];77(3):355–61. Available from: https://doi.org/10.1093/cid/ciad210

27. Pitzer VE, Chitwood M, Havumaki J, Menzies NA, Perniciaro S, Warren JL, et al. The Impact of Changes in Diagnostic Testing Practices on Estimates of COVID-19 Transmission in the United States. Am J Epidemiol [Internet]. 2021 Sep 1 [cited 2023 Oct 26];190(9):1908–17. Available from: https://doi.org/10.1093/aje/kwab089

28. Viboud C, Bjørnstad ON, Smith DL, Simonsen L, Miller MA, Grenfell BT. Synchrony, Waves, and Spatial Hierarchies in the Spread of Influenza. Science [Internet]. 2006 Apr 21 [cited 2024 Feb 29];312(5772):447–51. Available from: https://www.science.org/doi/10.1126/science.1125237

29. Eggo RM, Cauchemez S, Ferguson NM. Spatial dynamics of the 1918 influenza pandemic in England, Wales and the United States. J R Soc Interface [Internet]. 2011 Feb 6 [cited 2024 Dec 12];8(55):233–43. Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2010.0216

30. Gog JR, Ballesteros S, Viboud C, Simonsen L, Bjornstad ON, Shaman J, et al. Spatial Transmission of 2009 Pandemic Influenza in the US. PLOS Comput Biol [Internet]. 2014 Jun 12 [cited 2024 Dec 12];10(6):e1003635. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003635

Address for correspondence: Rafael Lopes, 350 George Street, New Haven, CT 06511.

Email: rafael.lopes@yale.edu Phone (+1) 475-300-8845

Nicole A. Swartwood, 90 Smith Street, Boston, MA 02120.

Email: nswartwood@hsph.harvard.edu. Phone (+1) 617-432-6171.

Figure 1: United States population and estimated cumulative SARS-CoV-2 infections per

capita distributed across the hexagonal grid. Panel **A, B,** and **C**: United States', New

England's and Connecticut's 2019 population Census estimates on the hexagonal grid. Panel **D,**

E, and **F**: United States', New England's, and Connecticut cumulative infections per 100,000

persons on the hexagonal grid (March 2020–December 2021).

Note: Numbers are given in a log 10 scale. Hexagons with no filling had no population reported

or infections counts ever estimated.

Figure 2: Estimated infections per capita of SARS-CoV-2 in the United States, March

2020–December 2021. Panel **A**: Time series of SARS-CoV-2 infection estimates for the United

States, the gray shaded areas show the first two large waves of infections. Panels **B, C, D**, and **E**:

Sequence of the spatially smoothed estimates of SARS-CoV-2 infections per capita associated

with Wave 1 at 4 time points. Panels **F, G, H**, and **I**: Sequence of the spatially smoothed

estimates of SARS-CoV-2 infections per capita associated with Wave 2 at 4 time points.

Figure 3: SARS-CoV-2 infections-per-capita surfaces progression and speed of invasion

for each wave. Comparison of SARS-CoV-2 infection Wave 1 and Wave 2 speed for 56- to 7-

days prior to each wave's infection peak. **See Figure S5 for a countor plot of waves**

progression.

Video. 'Wave1.mov' and 'Wave2.mov'

A

B

C

Population (log10 scale)

10^1 10^2 10^3 10^4 10^5 10^6 10^7

D

E

F

Cumulative Infections (log10 scale)
(March 2020 – December 2021)

10^1 10^2 10^3 10^4 10^5 10^6 10^7