Title: Quantifying the spatiotemporal dynamics of the first two epidemic waves of SARS-CoV-2 infections in the United States

Short Title: Dynamics of the first two epidemic waves of SARS-CoV-2 infections in the US

Rafael Lopes[1,2,*], 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], Nicole A. Swartwood[3, *,&]

[1] Department of Epidemiology of Microbial Diseases, Yale School of Public Health, Yale University, New Haven, Connecticut, United States of America

[2] Public Health Modeling Unit, Yale School of Public Health, Yale University, New Haven, Connecticut, United States of America

[3] Department of Global Health and Population, Harvard T.H. Chan School of Public Health, Harvard University, Boston, Massachusetts, United States of America

[4] Department of Health Policy, Stanford University, Stanford, California, United States of America

[5] Department of Biostatistics, Yale School of Public Health, Yale University, New Haven, Connecticut, United States of America

[&]Senior author

*rafael.lopes@yale.edu , nswartwood@hsph.harvard.edu

Abstract:

SARS-CoV-2 infection rates displayed strikingly organized patterns of temporal and spatial spread as new variants were introduced and subsequently transmitted within the United States. While these spatiotemporal "waves" of infection have been described previously, attempts to quantify the speed and extent of these waves have been limited. Here, we estimate and compare the wavefront speed and spatial expansion 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, we found that the second wave spread more rapidly than the first, possibly driven by multiple introduction events. These analyses highlight regional heterogeneity in epidemic dynamics and underscore the importance of localized public health measures in mitigating ongoing outbreaks.

Author Summary:

We developed tools to produce high-resolution visualizations and to quantify the complex patterns of spatial and temporal spread of SARS-CoV-2 during the first two major waves of infections in the contiguous United States.

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). A distinct feature of the subsequent SARS-CoV-2 epidemic, viewed from a national level, was the appearance of several periods of rapid increase and slower decline in estimated infection rates. These epidemic "waves" largely coincided with the introduction of genetically distinct SARS-CoV-2 variants (3,4) and were also affected by changes in the use and effectiveness of infection-mitigating interventions (5,6) and changes in population movement (7–10).

Here we quantify and compare the speed of wavefront expansion and the areal extent of wave spread over distinct epidemic waves. We apply a version of the Besag, York, and Mollié (BYM) spatial model (11,12) to produce temporally- and spatially-resolved estimates of SARS-CoV-2 infection rates (13) over the first two major waves of infections and use these estimates to estimate the speed of movement and spatial extent of wave involvement for the contiguous United States.

The two major epidemic waves we analyzed overlap with the introduction of the 'Alpha' (B.1.1.7) variant of concern (VOC) in mid-September 2020 and the 'Delta' (B.1.617.2) VOC in early July 2021 (14–18). We note, however, that the first wave we study begins to rise in advance of the widespread emergence of the Alpha VOC, hence we refer to these waves as Wave 1 and Wave 2, as they do not perfectly align with a singular genetic variant in circulation (17–20).

We note that, in addition to being associated with distinct viral variants, the two waves considered in this study were spreading in conditions that also differed in terms of levels of population immunity (conferred by either previous infections or vaccinations) and the types of non-pharmaceutical interventions being employed. Wave 1 occurred before vaccines were available, and when immunity was acquired as a result of infections with the wild-type variants (21,22). Wave 1 also began during a time of nationwide travel restrictions and widespread policies mandating mask use in congregate settings (5,7,9).

In contrast, wave 2 occurred when there were fewer mobility restrictions and fewer mandatory masking policies in place, though these policies varied markedly by geography within the US. Vaccines had also become available in January 2021 in the US, producing higher levels of apparent population immunity for wave 2 compared to wave 1 (21,22). We note wave 2 was largely attributable to the 'Delta' variant, which was characterized by its ability to escape immune responses acquired through either previous infection or vaccination.

Materials and 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, hospitalizations, and deaths, accounting for both under ascertainment and time lags in disease progression (13). Our dataset encompasses the period from March 2020 to December 2021. To estimate population denominators for our per 100,000 analyses, we used 2019 United States

population estimates via Meta, which provided population size estimates at a resolution of 30 square meters (23).

Regularization of geographical units

Counties were the smallest geographical unit with consistently available data for COVID-19 cases, hospitalizations, and deaths in the United States, which dictated the geographic resolution at which SARS-CoV-2 infections could be estimated. However, as the geographic size and population density of counties differ systematically across the US (larger area and sparser population density in the West than in the East), a county-level analysis would bias estimates of the wavefront speed and area of spread.

To improve our ability to detect spatial patterns in the data that were robust to administrative differences in county size, we first created a spatially regularized grid to which we allocated population and infections per 100,000 across the contiguous US. The grid comprises 7,517 unique hexagons, each enclosing 1,100 square kilometers. We employed an area-weighted approach to distribute population from the 30 square-meter estimates to the hexagons (**Fig 1A-1C**), assuming a constant population throughout the analysis period. Our daily estimates of SARS-CoV-2 infections per 100,000 for each hexagon were produced by the distribution of estimated infections from counties to hexagons (22) based on the fraction of each county's population contained in each hexagon, assuming that per 100,000 infection rates were distributed equally within a county. (**Fig 1D-1E**).

Assessment of spatial patterns of infections per 100,000

We fit a modified version of the BYM model as implemented in R-INLA (24), called BYM2, to estimate spatially smoothed rates of SARS-CoV-2 infections per 100,000 of population across the hexagonal grid. The BYM2 is a reparametrized version of the original Besag (12) model, which improves the assignment of prior distributions for the corresponding model parameters and the subsequent interpretation of parameters (11). Specifically, our model is given as

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

where $Z(A_i)$ the infections per 100,000 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 priors used, see the **Supplementary Material.** We fit the equation 1 model separately for each day over March 2020 to December 2021. We produced a sequence of these daily spatial distributions, which we then used to calculate the speed of wave front movement and spatial 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.

Definition of a wave

To formally determine hexagons within a wave, we defined a threshold of 190 SARS-CoV-2 infections per 100,000 persons or more, which we coded as being part of an infection wave. Any hexagon within a wave could leave the wave if infections fell below that threshold. For our main analysis, we used the above-mentioned threshold of 190 daily infections per 100,000 persons to characterize a hexagon as being within a wave. This value represents the 75th percentile of the infections-per-100,000 value distribution during the study period (**S1 Fig**). We conducted

sensitivity analyses using 50th (127 infections per 100,000) and 90th (233 infections per 100,000) percentiles of infections per 100,000 persons as alternative threshold values (**S2 Fig)**. To allow for multiple infection centers, we did not limit our definition to a single contiguous set of hexagons.

Using this definition, we categorized Wave 1 as the sets of contiguous hexagons above the threshold within the period ranging from September 17, 2020 to February 19, 2021, and Wave 2 as the sets of contiguous hexagons above the 190 infections per 100,000 threshold during the period from July 3, 2021 to September 4, 2021. We identified the epicenters of each wave by visual inspection.

Calculation of wavefront speed and areal wave expansion

We first identified the leading edge, or wavefront, of each wave. This wavefront was defined as the two-dimensional boundary between hexagons included and excluded from a wave on each day of the analysis. After delineating the wavefront, we estimated two quantities: the daily areal wave expansion and the daily wavefront speed. To calculate the areal wave expansion rate, we tracked the daily change in the number of hexagons included in a wave and multiplied this by the area of each hexagon (1,100 km^2) to calculate the expansion rate in km^2. This quantity is the difference between the area of the wave at day $d+1$ and day d, giving an areal wave expansion rate measured in km^2/day.

To estimate the wavefront speed, we identified the vertices of the line-segments comprising the wavefront on each day, *d.* We calculated the distance between each vertex and the nearest point

152 on the subsequent wavefront at day *d+1*, which provides a measure of distance travelled per day
153 (km/day), or speed for each vertex of the wavefront. This allows us to capture the heterogeneity
154 in speed along the length of the wavefront. For each day, we estimated the daily mean speed as
155 the mean of all wave vertex speeds estimated for that day. We also summarized the speed of each
156 wave by calculating the mean of the estimated speed for all days and wave vertices with equal
157 weights. Finally, we estimated the length of the wavefront that was moving at a rate greater than
158 the median of the daily maximum estimated speeds. This length was calculated as the sum of the
159 line-segment lengths adjacent to a vertex with an estimated speed greater than the median daily
160 maximum speed.
161
162 **Statistical analysis**
163 All analyses were performed using R Statistical Software (v4.3.0) (25). We used the **INLA**, **sf**,
164 **sp**, **spdep**, **areal**, and **magick** R packages (24,26–30). All the maps were generated using
165 United States, state, and county borderlines maps in the public domain from the Census Bureau,
166 which were downloaded through the R package Tigris (31,32).
167
168 **Results**
169
170 **Spatiotemporal patterns of SARS-CoV-2 infections across the United States**
171 As described in the **Materials and Methods**, we categorized infections as belonging to Wave 1
172 (September 17, 2020 to February 19, 2021) or Wave 2 (July 3, 2021 to September 4, 2021) (**Fig**
173 **2A**). **Fig 2B-2I** shows the smoothed estimates of SARS-CoV-2 infections per 100,000 from the
174 BYM2 model on eight dates, leading up to the peak of each of the two waves analyzed. Wave 1

originated with a set of hexes spanning central South Dakota, eastern North Dakota, and northeastern Montana (**Fig 2B**) in mid-September of 2020. Then the wave extended south and to both coasts, achieving a peak of over 1.5 million prevalent infections/day by mid-November 2020 (**Fig 2A and 2C-2E**). Wave 2 originated in the Ozarks (southern Missouri and northern Arkansas) (**Fig 2F**) in early July 2021 and then expanded further south (**Fig 2G-2I**). Secondary epicenters of infection appeared later in July in the Pacific Northwest (**Fig 2G**), and the infection wave subsequently spread throughout the western United States (**Fig 2G-2I**). This wave had a peak of 300 per 100,000 infections/day by early September 2021.

Animations of the infection waves are provided in S**1 Movie [link to: Movie1.mov]** and S**2 Movie [link to: Movie2.mov],** which present maps of the daily, spatially smoothed estimated infections for waves 1 and 2, respectively. All maps were generated using United States, state, and county borderlines maps in the public domain from the Census Bureau, which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infection estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

Wave expansion

Fig 2B-2C shows contour plots of the speed of expansion for each wave. Wave 1 expanded from an area of 134,200 km^2 to 6,515,300 km^2 between September 8, 2020 and November 11, 2020 (**Fig 3, '1st Wave' bars**). Wave 2 expanded from an area of 23,100 km^2 to 7,573,500 km^2 between July 7, 2021 and September 4, 2021 (**Fig 3, '2nd Wave' bars**). The wave-like patterns were robust to the choice of threshold (**S2** and **S3 Figs**).

Fig 3 compares the areal expansion of each wave for each analyzed day before each wave's respective peak. Wave 1 had a mean areal expansion rate of 101,287 km^2/day (median= 64,900 interquartile range (IQR): 20,350–199,650 km^2/day). Wave 2 had a slightly higher mean areal expansion rate of 119,848 km^2/day (median = 69,300; IQR = 29,700–192,500 km^2/day) during the analysis period.

Wave speed

The overall mean wavefront speed of Wave 1 was 20.3 km/day (median = 20.6; IQR = 0–35.6 km/day) across all geographies and the 63 days of analysis. The daily mean wavefront speed had an IQR of 4.4–33.4 km/day. Wave 2 had a higher overall mean wavefront speed of 25.9 km/day (median = 20.6; IQR = 0–35.6 km/day). The daily mean wavefront had an IQR of 7.0–32.9 km/day. **S1 Table** provides daily mean and median wavefront speeds for each wave.

There was heterogeneity in the estimated daily wavefront speed across the edge of each wave. **Fig 4** shows the speed of each wave at four dates leading up to each wave's infection peak. To understand changes in speed over the course of each wave, we calculated mean speed over three time periods: 63–43 days, 42–22 days, and 21–0 days prior to each wave's peak. These speed measures increased over the course of each wave, and for each time period, Wave 2 was faster than Wave 1, with mean speeds of 3.8 (median = 0; IQR = 0–0), 17.0 (median = 20.6; IQR = 0–20.6), and 33.9 km/day (median = 35.6; IQR = 20.6–54.4) and 6.8 (median = 0; IQR = 0–0), 19.8 (median = 20.6; IQR = 0-35.6) , and 40.8 (median = 35.6; IQR = 20.6-54.4) km/day, respectively.

Because wavefronts varied in length, the percentage of the wavefront at maximal speed changed over time; to quantify this change, we calculated the length of the wavefront exceeding the median maximum speed over the study period. At its longest, the wavefront of Wave 1 had 7,017 kilometers moving at a speed greater than or equal to the overall median maximal speed at 18 days before its infections per 100,000 peak. In comparison, Wave 2 had a maximal wavefront of 9,157 kilometers at 9 days before its peak. The daily lengths of the wavefront moving faster than the median of the estimated daily maximum speeds are in **S1 Table**.

Animations of daily wavefront speeds are provided in S**3 Movie [link to file: Movie3.mov]** and **S4 Movie [link to file: Movie4.mov],** for waves 1 and 2, respectively. All maps were generated using United States, state, and county borderlines maps in the public domain from the Census Bureau, which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infection estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

Discussion

In this study, we quantified the wavefront speed and the areal expansion rate of the first two major SARS-CoV-2 epidemic waves in the contiguous United States. By applying a Bayesian spatial modeling framework to spatially resolved infection estimates, we were able to provide new visualizations of wave movement and quantify and compare wave behavior. To produce these estimates, we adopted a definition of the wave front boundary corresponding to >190 infections per 100,000 per day. While this produces a wavefront that lags behind the first

reported case at any given location, it provides a more stable measure of epidemic spread that corresponds to meaningful levels of community transmission.

Our findings suggest that the first and second waves differed in both their origin and expansion patterns. Wave 1 originated in parts of northeastern Montana, eastern North Dakota, and central South Dakota in September 2020 (**Fig 2B-2E** and S**2 Movie**). The estimated point of origin coincides with a motorcycle rally that brought over 460,000 attendees to Sturgis, South Dakota, in August 2020 and was later epidemiologically linked to elevated local COVID-19 rates and to the interstate spread of infection (33,34).

In contrast, Wave 2 had multiple epicenters, beginning in the Ozarks (Fig 2F and S2 Movie) during a time when indoor music venues were reopening, and local masking policies were relaxed (35,36). Shortly after, secondary high-incidence epicenters appeared in the Pacific Northwest. The decentralized and near-simultaneous emergence of multiple hotspots during Wave 2 aligns with earlier work demonstrating that both domestic and international travel networks have facilitated the rapid dispersal of viruses (6,37,38). Wave 2 coincided with the emergence and national spread of the Delta variant (B.1.617.2), which became dominant in the United States during this period (18).

While we found that Wave 2 had both a higher maximal speed of wavefront movement and achieved its most significant areal expansion more quickly than Wave 1, the similarity in estimates of wavefront movement speed and the growth rate in wave size for these first two waves is striking. This similarity is remarkable considering the distinct virological,

267 immunological, and policy contexts in which each wave occurred. Wave 1 was due to wildtype
268 SARS-CoV-2 as well as the 'Alpha' variant (37,38). Wave 2 was driven by the 'Delta' variant
269 (17,18,35) and emerged at a time when approximately 50-60% of the US adult population had
270 received at least one COVID-19 vaccination dose (39,40). Regional and secular differences in
271 non-pharmaceutical intervention policies, and differences in the levels and types of infection-
272 and vaccine-induced population immunity at the time of these two waves (17), make the
273 similarity between the movement patterns of these waves particularly surprising. In a previous
274 analysis, we estimated state-level immunological protection against SARS-CoV-2 infection
275 among US adults. At the onset of Wave 1 in September 2020, we estimated that ~5% of adults
276 were protected against infection. This protection varied minimally between US regions, ranging
277 from ~4% in the Midwest to ~7% in the Northeast. This immunological environment contrasted
278 with that of Wave 2; in July 2021, an estimated ~34% of the US population was protected
279 against SARS-CoV-2 infection, ranging from ~31% in the South to ~38% in the Northeast (41).
280 The consistent speed of expansion of the two waves, each in a different immunological context,
281 may suggest that these expansion rates are relatively stable across immunological contexts.
282 Alternatively, it could be that different factors (population immunity, application of non-
283 pharmaceutical interventions, pathogen characteristics, and underlying patterns of human
284 mobility) had approximately offsetting effects in the two waves we examined (7,42).
285
286 Our effort to characterize and provide quantitative estimates of the speed of spread to epidemic
287 waves required several simplifying assumptions. Importantly, we elected to use estimates of
288 infections from a nowcasting model (rather than case notification data) as our input data. As
289 diagnosis and case reporting varied markedly in quality and completeness over the epidemic

(43), this choice allowed for more consistent input data. Another challenge we faced was systematic variability in county geography and population density across the contiguous United States. We also chose to distribute estimated infections on a hexagonal grid before estimating wavefront speed and areal patterns of wave expansion to address potential bias associated with the irregular shapes and sizes of US counties (the smallest unit at which infection estimates were possible). Nevertheless, we could have introduced bias, since cases within counties were not randomly distributed; however, this effect should be modest given the small sizes and large numbers of hexagons. Finally, we arbitrarily selected the threshold of infections per 100,000 to define wave membership but note that our sensitivity analyses showed that the estimated metrics describing wavefront speed and areal expansion rate were similar across alternative thresholds.

Factors that affect the diffusion and speed of epidemic waves are an area of substantial interest for other viral diseases (44), like influenza (45–47), Ebola (e.g.,(48)), HIV (49), measles (50), and parasitic diseases such as malaria (51). While our work on SARS-CoV-2 does not 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 wavefront speed and rate of areal expansion of epidemic waves, such as those we employed here, is a necessary first step. Related work has evaluated the quantification of wave dynamics for point pattern disease data (52–54)and from trend surface analysis (55–57). Further research to explore how the pathogen, environment, and host characteristics affect the speed and pattern of epidemic expansion may improve epidemic predictions and assist in the planning and deployment of spatially defined interventions.

Funding

This project is supported by Cooperative Agreement NU38OT000297 from the Centers for Disease Control and Prevention (CDC) and the Council of State and Territorial Epidemiologists (CSTE) (NAS, FK, JAS, TC and NAM), SHEPheRD Contract 200-2016-91779 from the CDC, and the CDC Broad Agency Announcement Contract 75D30122C14697 (RL, YL, MHC, NDG, JLW, and TC). This work does not necessarily represent the views of the CDC or CSTE. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Data availability:

Daily county estimates of per 100,000 infections used in the analysis are available at Harvard Dataverse (58). We have made our repository, https://github.com/covidestim/waves, with the relevant scripts and necessary origination data, such as hexagonally distributed infections per 100,000, publicly available. These materials allow the reproduction of all the figs and analytical products in this analysis. We will also make a static repository of all codes and intermediary data products available on Zenodo: https://zenodo.org/records/18624632.

Author contribution – CRediT Taxonomy

Conceptualization: RL, NAS, TC

Methodology: RL, NAS, JLW, TC

Investigation: RL, NAS

Visualization: RL, NAS, YL

Funding acquisition: NAM, JAS, NDG, TC

Supervision: NAS, JAS, NAM, JLW, NDG, TC

Writing – original draft: RL, NAS

Writing – review & editing: all authors

Competing interests

I have read the journal's policy, and the authors of this manuscript have the following competing interests: NDG is a paid consultant for BioNTech. All the other authors have declared that no competing interests exist.

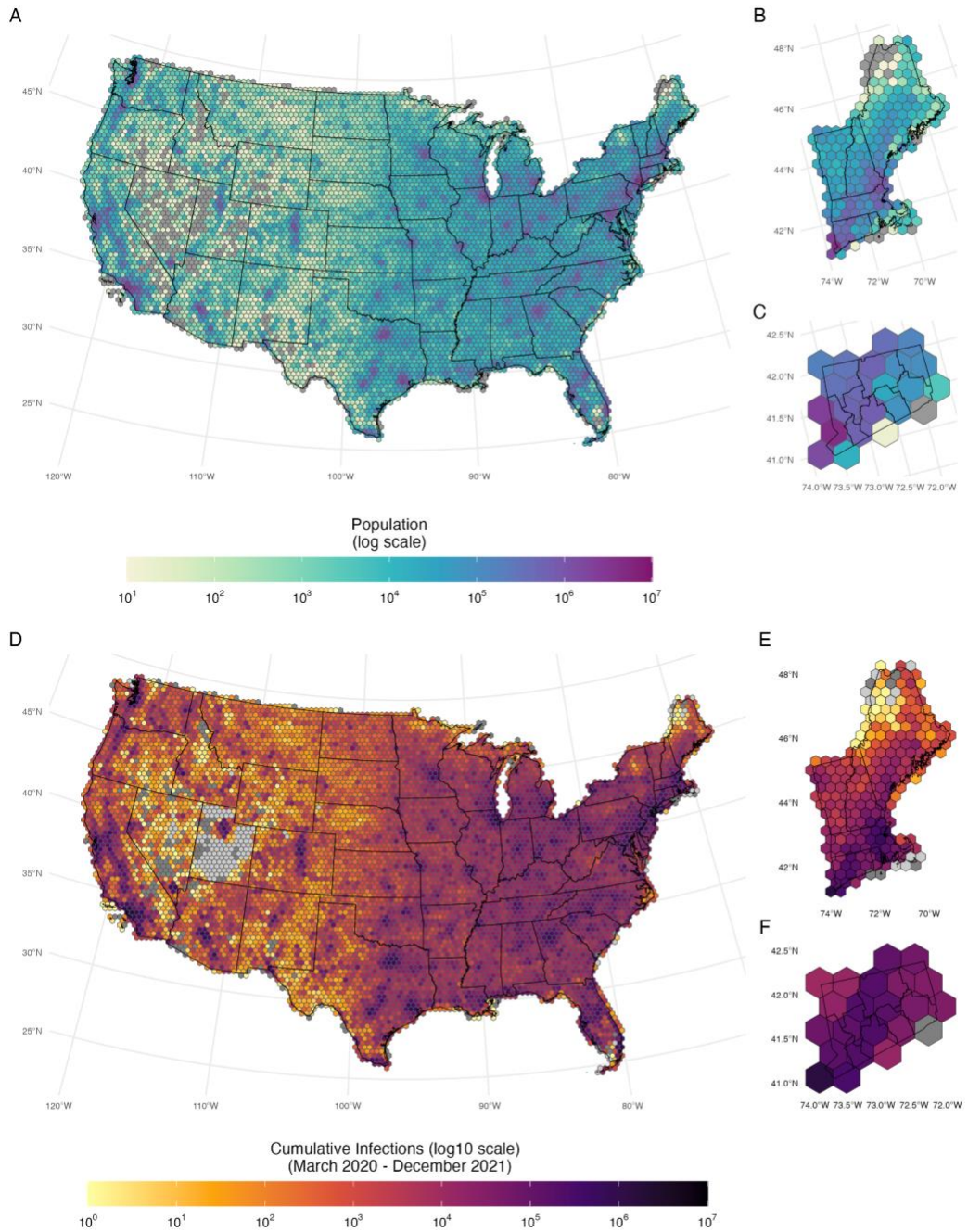

Fig 1: United States population and estimated cumulative SARS-CoV-2 infections per 100,000 distributed across the hexagonal grid. Panel **A, B,** and **C**: United States', New England's, and Connecticut's Meta 30m population estimates on the hexagonal grid. Panel **D, E,** and **F**: United States, New England, and Connecticut cumulative infections per 100,000 persons

on the hexagonal grid (March 2020–December 2021). Grey hexagons indicate no population (panels **A** to **C)** and no infections (panels **D** to **F),** void hexagons indicate no infections ever estimated (panels **D** to **F**)**.** All maps were generated using United States, state, and county borderlines maps in the public domain from the Census Bureau, which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infection estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

Note: Numbers are given in a log 10 scale.

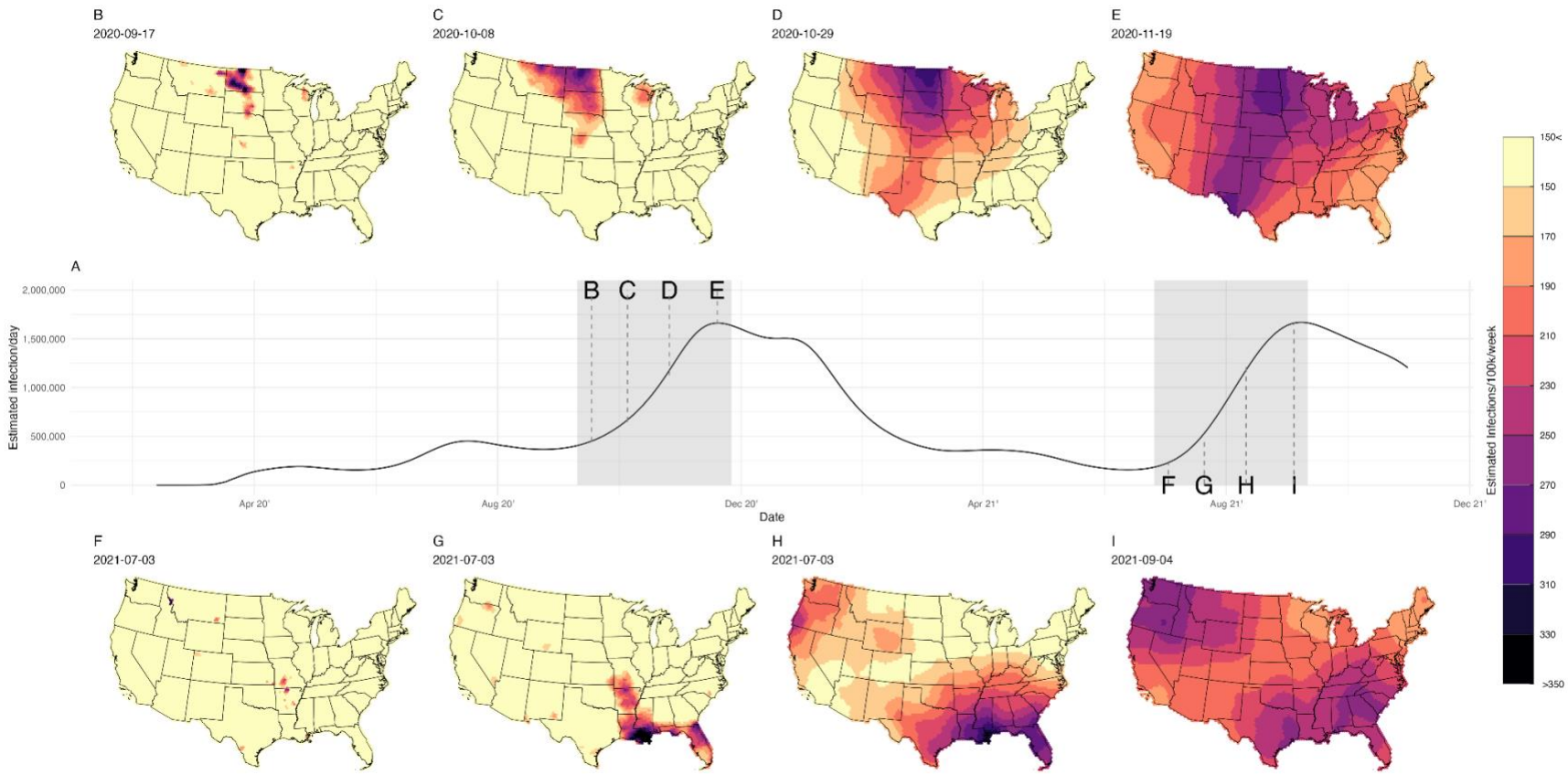

Fig 2: Estimated infections per 100,000 of SARS-CoV-2 in the United States, March 2020–December 2021. Panel **A**: Time series of SARS-CoV-2 infection estimates for the United States; the gray shaded areas show the first two large waves of infections. Panels **B, C, D**, and **E**: Sequence of the spatially smoothed estimates of SARS-CoV-2 infections per 100,000 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 100,000 associated with Wave 2 at 4 time points. All maps were generated using United States, state, and county borderlines maps in the public domain

from the Census Bureau, which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infection estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products. For a more detailed visualization of each wave, see the movies in the online supplemental material

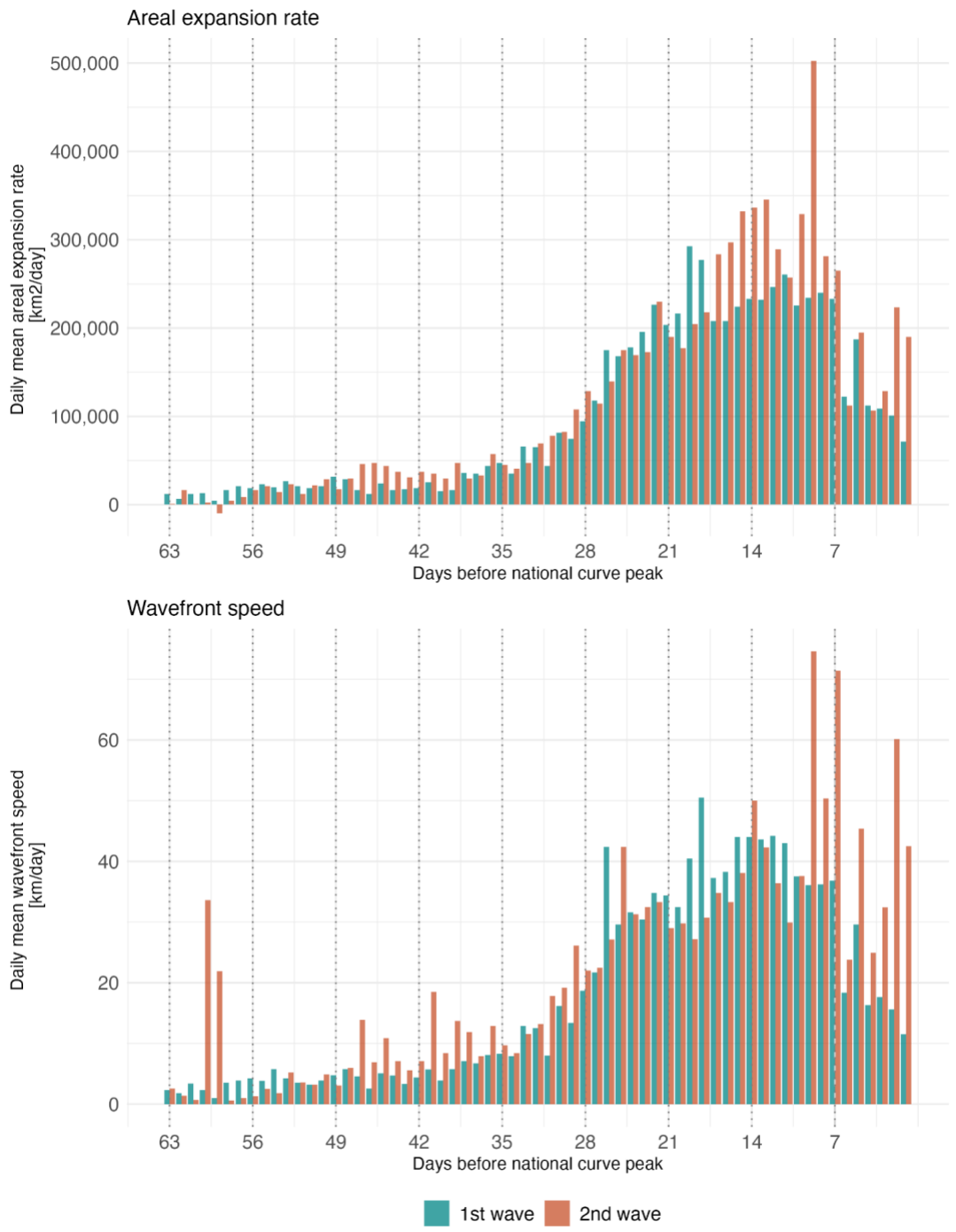

Fig 3: Daily mean areal expansion rate and wavefront speed for SARS-CoV-2 infection waves for 63 days preceding the infection per capita peak. (Wave 1: September 8, 2020–November 11,

2020; Wave 2: July 7, 2021–September 4, 2021). **Top Panel:** Mean daily areal expansion rate. **Bottom Panel:** Mean daily wavefront speed.

Note: Areal expansion rate is the daily change in wave area. Wavefront speed is the distance from a vertex of the hexagonal grid at day d to the nearest point on the line at day d+1.

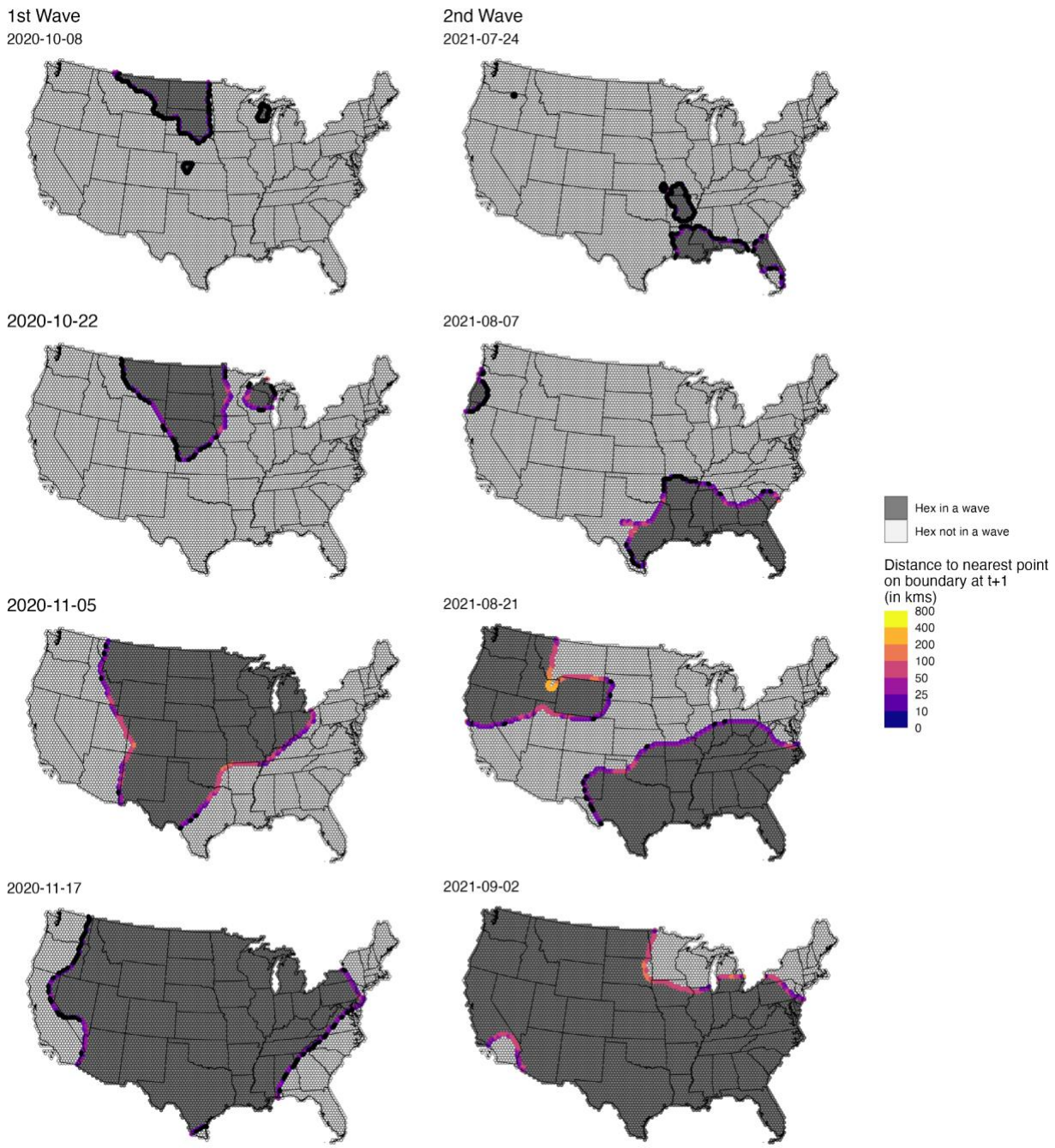

Fig 4: Progression of the wave boundary measured as the distance to the nearest point on the boundary at t+1 (in km/day). Each panel shows a specific date; the ones in the right column are related to the first wave and the ones in the left column to the second wave. Vertices

in black show points with zero speed (i.e., their location is the same for t and t+1. All maps were generated using United States, state, and county borderlines maps in the public domain from the Census Bureau, which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infection estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

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. Holmes EC. The Emergence and Evolution of SARS-CoV-2. Annu Rev Virol [Internet]. 2024 Apr 17 [cited 2024 May 1]; Available from: https://www.annualreviews.org/content/journals/10.1146/annurev-virology-093022-013037

4. Martin DP, Weaver S, Tegally H, San JE, Shank SD, Wilkinson E, et al. The emergence and ongoing convergent evolution of the SARS-CoV-2 N501Y lineages. Cell [Internet]. 2021 Sept 30 [cited 2023 Oct 17];184(20):5189-5200.e7. Available from: https://www.sciencedirect.com/science/article/pii/S0092867421010503

5. 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

6. Davis JT, Chinazzi M, Perra N, Mu K, Pastore y Piontti A, Ajelli M, et al. Cryptic transmission of SARS-CoV-2 and the first COVID-19 wave. Nature [Internet]. 2021 Dec [cited 2024 Oct 16];600(7887):127–32. Available from: https://www.nature.com/articles/s41586-021-04130-w

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 July 8 [cited 2024 Oct 2];7(1):230. Available from: https://www.nature.com/articles/s41597-020-00575-2

9. Kishore N, Kahn R, Martinez PP, De Salazar PM, Mahmud AS, Buckee CO. Lockdowns result in changes in human mobility which may impact the epidemiologic dynamics of SARS-CoV-2. Sci Rep [Internet]. 2021 Mar 26 [cited 2023 Dec 19];11(1):6995. Available from: https://www.nature.com/articles/s41598-021-86297-w

10. Harris JE. Mobility was a significant determinant of reported COVID-19 incidence during the Omicron Surge in the most populous U.S. Counties. BMC Infect Dis [Internet]. 2022 Aug 15 [cited 2025 June 30];22:691. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9376582/

11.Riebler A, Sørbye SH, Simpson D, Rue H. An intuitive Bayesian spatial model for disease mapping that accounts for scaling [Internet]. arXiv; 2016 [cited 2024 Oct 2]. Available from: http://arxiv.org/abs/1601.01180

12.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

13.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

14.Petrone ME, Rothman JE, Breban MI, Ott IM, Russell A, Lasek-Nesselquist E, et al. Combining genomic and epidemiological data to compare the transmissibility of SARS-CoV-2 variants Alpha and Iota. Commun Biol [Internet]. 2022 May 11 [cited 2023 June 14];5(1):1–10. Available from: https://www.nature.com/articles/s42003-022-03347-3

15.Faucher B, Sabbatini CE, Czuppon P, Kraemer MUG, Lemey P, Colizza V, et al. Drivers and impact of the early silent invasion of SARS-CoV-2 Alpha. Nat Commun [Internet]. 2024 Mar 9 [cited 2024 Mar 12];15(1):2152. Available from: https://www.nature.com/articles/s41467-024-46345-1

16.Brunner FS, Brown MR, Bassano I, Denise H, Khalifa MS, Wade MJ, et al. City-wide wastewater genomic surveillance through the successive emergence of SARS-CoV-2 Alpha and Delta variants. Water Res [Internet]. 2022 Nov 1 [cited 2023 Oct 22];226:119306. Available from: https://www.sciencedirect.com/science/article/pii/S0043135422012519

17.Earnest R, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, et al. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA. Cell Rep Med [Internet]. 2022 Apr 19 [cited 2023 June 14];3(4). Available from: https://www.cell.com/cell-reports-medicine/abstract/S2666-3791(22)00090-8

18.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

19.Carabelli AM, Peacock TP, Thorne LG, Harvey WT, Hughes J, de Silva TI, et al. SARS-CoV-2 variant biology: immune escape, transmission and fitness. Nat Rev Microbiol [Internet]. 2023 Mar [cited 2023 Oct 17];21(3):162–77. Available from: https://www.nature.com/articles/s41579-022-00841-7

20.Figgins MD, Bedford T. SARS-CoV-2 variant dynamics across US states show consistent differences in effective reproduction numbers [Internet]. medRxiv; 2022 [cited 2024 May 1]. p. 2021.12.09.21267544. Available from: https://www.medrxiv.org/content/10.1101/2021.12.09.21267544v2

21.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

22.Klaassen F, Swartwood NA, Chitwood MH, Lopes R, Haraguchi M, Salomon JA, et al. National- and state-level SARS-CoV-2 immunity trends from January 2020 to December 2023: a mathematical modeling analysis [Internet]. medRxiv; 2024 [cited 2025 Sept 11]. p. 2024.10.22.24315935. Available from: https://www.medrxiv.org/content/10.1101/2024.10.22.24315935v1

23.United States: High Resolution Population Density Maps + Demographic Estimates | Humanitarian Dataset | HDX [Internet]. [cited 2025 May 23]. Available from: https://data.humdata.org/dataset/united-states-high-resolution-population-density-maps-demographic-estimates

24.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

25.R Core Team. R: A Language and Environment for Statistical Computing_. R Foundation for Statistical Computing, Vienna, Austria. 2024.

26.Pebesma E. Simple Features for R: Standardized Support for Spatial Vector Data. R J [Internet]. 2018 [cited 2024 Oct 4];10(1):439. Available from: https://journal.r-project.org/archive/2018/RJ-2018-009/index.html

27.Pebesma EJ, Bivand RS. Classes and methods for spatial data in R.

28.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

29.Prener CG, Revord CK. areal: An R package for areal weighted interpolation. J Open Source Softw [Internet]. 2019 May 19 [cited 2024 Oct 4];4(37):1221. Available from: https://joss.theoj.org/papers/10.21105/joss.01221

30.magick: Advanced Graphics and Image-Processing in R [Internet]. [cited 2024 Oct 4]. Available from: https://ropensci.r-universe.dev/magick

31.Bureau UC. Census.gov. [cited 2025 May 26]. TIGER/Line Shapefiles. Available from: https://www.census.gov/geographies/mapping-files/time-series/geo/tiger-line-file.html

32.Walker K, Rudis B. tigris: Load Census TIGER/Line Shapefiles [Internet]. 2025 [cited 2025 May 26]. Available from: https://cran.r-project.org/web/packages/tigris/index.html

33.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.

34.Mervosh S, Lu D, Swales V. See Which States and Cities Have Told Residents to Stay at Home. The New York Times [Internet]. 2020 Mar 24 [cited 2025 May 12]; Available from: https://www.nytimes.com/interactive/2020/us/coronavirus-stay-at-home-order.html

35.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/

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

37.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

38.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

39.Diesel J. COVID-19 Vaccination Coverage Among Adults — United States, December 14, 2020–May 22, 2021. MMWR Morb Mortal Wkly Rep [Internet]. 2021 [cited 2026 Jan 23];70. Available from: https://www.cdc.gov/mmwr/volumes/70/wr/mm7025e1.htm

40.Datar RS, Fette LM, Hinkelman AN, Hammershaimb EA, Friedman-Klabanoff DJ, Mongraw-Chaffin M, et al. Factors associated with COVID-19 vaccination during June–October 2021: A multi-site prospective study. Vaccine [Internet]. 2023 May 11 [cited 2026 Jan 22];41(20):3204–14. Available from: https://pmc.ncbi.nlm.nih.gov/articles/PMC10063571/

41.Klaassen F, Swartwood NA, Chitwood MH, Lopes R, Haraguchi M, Salomon JA, et al. National- and State-level SARS-CoV-2 Immunity Trends From January 2020 to December 2023: a Mathematical Modeling Analysis. J Infect Dis [Internet]. 2025 Oct 23 [cited 2026 Jan 23];jiaf532. Available from: https://doi.org/10.1093/infdis/jiaf532

42.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 21];69. Available from: https://www.cdc.gov/mmwr/volumes/69/wr/mm6947e1.htm

43.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 Sept 1 [cited 2023 Oct 26];190(9):1908–17. Available from: https://doi.org/10.1093/aje/kwab089

44.Dellicour S, Bastide P, Rocu P, Fargette D, Hardy OJ, Suchard MA, et al. How fast are viruses spreading in the wild? Dushoff J, editor. PLOS Biol [Internet]. 2024 Dec 3 [cited 2025 Mar 28];22(12):e3002914. Available from: https://dx.plos.org/10.1371/journal.pbio.3002914

45.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

46.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

47.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 June 12 [cited 2024 Dec 12];10(6):e1003635. Available from: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1003635

48.Wesolowski A, Buckee CO, Bengtsson L, Wetter E, Lu X, Tatem AJ. Commentary: Containing the Ebola Outbreak - the Potential and Challenge of Mobile Network Data. PLoS Curr [Internet]. 2014 Sept 29 [cited 2025 May 11];6:ecurrents.outbreaks.0177e7fcf52217b8b634376e2f3efc5e. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4205120/

49.Faria NR, Rambaut A, Suchard MA, Baele G, Bedford T, Ward MJ, et al. The early spread and epidemic ignition of HIV-1 in human populations. Science [Internet]. 2014 Oct 3 [cited 2024 Oct 13];346(6205):56–61. Available from: https://www.science.org/doi/10.1126/science.1256739

50.Grenfell BT, Bjørnstad ON, Kappey J. Travelling waves and spatial hierarchies in measles epidemics. Nature [Internet]. 2001 Dec 13 [cited 2024 Feb 3];414(6865):716–23. Available from: https://www.nature.com/articles/414716a

51.Wesolowski A, Eagle N, Tatem AJ, Smith DL, Noor AM, Snow RW, et al. Quantifying the Impact of Human Mobility on Malaria. Science [Internet]. 2012 Oct 12 [cited 2025 May 12];338(6104):267–70. Available from: https://www.science.org/doi/10.1126/science.1223467

52.Schliep EM, Gelfand AE. Velocities for spatio-temporal point patterns. Spat Stat [Internet]. 2019 Mar 1 [cited 2025 Mar 17];29:204–25. Available from: https://www.sciencedirect.com/science/article/pii/S2211675318301714

53.Avellaneda FR, Mateu J, Moraga P. Estimating velocities of infectious disease spread through spatio-temporal log-Gaussian Cox point processes [Internet]. arXiv; 2024 [cited 2025 Mar 17]. Available from: http://arxiv.org/abs/2409.05036

54.Diggle P, Zheng P, Durr P. Nonparametric estimation of spatial segregation in a multivariate point process: bovine tuberculosis in Cornwall, UK. J R Stat Soc Ser C Appl Stat [Internet]. 2005 [cited 2025 Jan 8];54(3):645–58. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9876.2005.05373.x

55.Pioz M, Guis H, Calavas D, Durand B, Abrial D, Ducrot C. Estimating front-wave velocity of infectious diseases: a simple, efficient method applied to bluetongue. Vet Res [Internet]. 2011 [cited 2024 Apr 4];42(1):60. Available from: http://www.veterinaryresearch.org/content/42/1/60

56.Freire CC de M, Iamarino A, Neto DF de L, Sall AA, Zanotto PM de A. Spread of the pandemic Zika virus lineage is associated with NS1 codon usage adaptation in humans [Internet]. bioRxiv; 2015 [cited 2024 Sept 22]. p. 032839. Available from: https://www.biorxiv.org/content/10.1101/032839v1

57.Tisseuil C, Gryspeirt A, Lancelot R, Pioz M, Liebhold A, Gilbert M. Evaluating methods to quantify spatial variation in the velocity of biological invasions. Ecography [Internet]. 2016 [cited 2025 Apr 1];39(5):409–18. Available from: https://onlinelibrary.wiley.com/doi/abs/10.1111/ecog.01393

58.Klaassen F. covidestim updated: Estimating SARS-CoV-2 infections and immunity over the entire pandemic [Internet]. Harvard Dataverse; 2024 [cited 2025 Nov 25]. Available from: https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/G2ZXJG

List of legends - Supporting Information files

S1 Movie.mov - Wave 1 spreading sequential maps

S2 Movie.mov - Wave 2 spreading sequential maps

S3 Movie.mov - Wave 1 wavefront speed sequential maps

S4 Movie.mov - Wave 2 wavefront speed sequential maps

Supporting Information

S1_Appendix.docx

S1 Fig – Empirical cumulative density function (ECDF) of the risk surface values. The ECDF shows that a threshold of 190 infections per 100,000 is indicated by the vertical dashed line at which the ECDF crosses the 75th percentile. The ECDF for the lower and upper bounds is shown in grey.

S2 Fig – Sensitivity analysis of areal wave expansion to the threshold values. Areal wave expansion (km2/day) for different thresholds of infection per 100,00 in the progression calculation of the surfaces. Panel **A** is built with a threshold of 127 or more infections per 100,000; panel **B** is built with a threshold of 233 infections per 100,000. As in **Fig 3C**, we observe a maximal speed and a steep decrease after the peak, and the second wave had a higher invasion speed and encompassed a larger area at peak than the first wave.

S3 Fig – Infection per 100,000 surface on a continuous scale of values. As expected, the wave-like pattern holds independently of the scale to be displayed, and as being an output of spatial smooth model, the continuous scale gives a less defined border to the risk surface expansion. All maps were generated using United States, state, and county borderlines maps in public domain from the Census Bureau which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infections estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

S4 Fig - Infection per 100,000 with a threshold equal to the mean of the risk values distribution (127 infections per 100,000). With a lower threshold showing on the map, the spread process seems to happen faster. All maps were generated using United States, state, and county borderlines maps in public domain from the Census Bureau which were downloaded through the R package Tigris (30). The shapefile generated for this analysis with the population estimates and cumulative infections estimates can be found at: https://github.com/covidestim/waves/tree/waves-manuscript/Data/data-products.

S1 Table – Daily estimates of wavelength, areal growth, and speed for each wave