Supplementary Materials of: "Estimating the establishment of local transmission and the cryptic phase of the COVID-19 pandemic in the USA"

Jessica T. Davis[1*], Matteo Chinazzi [1*], Nicola Perra [2*,1], Kunpeng Mu[1], Ana Pastore y Piontti[1], Marco Ajelli[3,4], Natalie E. Dean[5], Corrado Gioannini[6], Maria Litvinova[6], Stefano Merler[3], Luca Rossi[6], Kaiyuan Sun[7], Xinyue Xiong[1], M. Elizabeth Halloran[8,9], Ira M. Longini Jr.[5], Cécile Viboud[7], Alessandro Vespignani[1,6,†]

[1]Laboratory for the Modeling of Biological and Socio-technical Systems, Northeastern University, Boston, MA USA
[2]Networks and Urban Systems Centre, University of Greenwich, London, UK
[3]Bruno Kessler Foundation, Trento Italy
[4]Department of Epidemiology and Biostatistics, Indiana University School of Public Health, Bloomington, IN, USA
[5]Department of Biostatistics, College of Public Health and Health Professions, University of Florida, Gainesville, USA
[6]ISI Foundation, Turin, Italy
[7]Fogarty International Center, NIH, USA
[8]Fred Hutchinson Cancer Research Center, Seattle, WA, USA
[9]Department of Biostatistics, University of Washington, Seattle, WA. USA
[*]These authors contribute equally to this work
[†]To whom correspondence should be addressed; E-mail: a.vespignani@northeastern.edu.

Here, we provide the details about the model calibration, present the sensitivity analysis of some key parameters, and describe the details of the importation sources estimation. We also include an analysis of the empirical data, several indicators (air traffic, population, density), and the data obtained via the model.

Global Epidemic and Mobility Model

We adopt the Global Epidemic and Mobility model (GLEAM), a stochastic, spatial, epidemic model based on a metapopulation approach that has been used and published previously (*1, 2*). In the model, the world is divided into over 3,200 geographic subpopulations constructed using a Voronoi tessellation of the Earth's surface. Subpopulations, centered around major

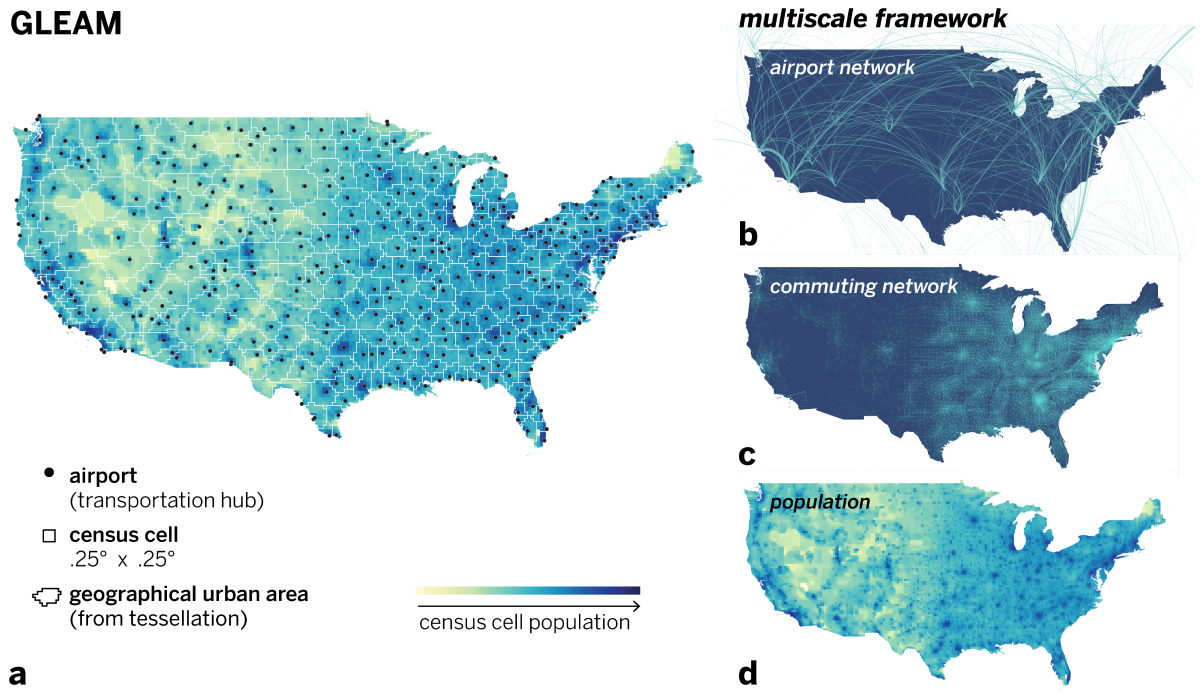

Figure 1: Schematic representation of GLEAM. (A) The subpopulation structure of the US. Subpopulations are geographic regions, formed from the Voronoi tessellation, that are centered around airports. They are comprised of census cells that are approximately 25km x 25km. (B) Diagram of the origin destination airport network (long range mobility network) in the US. (C) Diagram of the commuting network (short-range mobility network) in the US. (C) The population layer of GLEAM showing the population size of census cells.

transportation hubs (e.g. airports), consist of cells with a resolution of 15 x 15 arc minutes (approximately 25 x 25 kilometers). High resolution data are used to define the population of each cell (*3*). Other attributes of individual subpopulations, such as age specific contact patterns, health infrastructure, etc., are added according to available data (*4*).

GLEAM integrates a human mobility layer, represented as a network, using both short-range (i.e. commuting) and long-range (i.e. flights) mobility data from the Offices of Statistics for 30 countries on 5 continents as well as the Official Aviation Guide (OAG) and IATA databases (updated in 2019) (*5, 6*). The air travel network consists of the daily passenger flows between airport pairs (origin and destination) worldwide mapped to the corresponding subpopulations. We define a worldwide homogeneous standard for GLEAM to overcome differences in the spatial resolution of the commuting data across different countries. Where information is not available, the short-range mobility layer is generated synthetically by relying on the "gravity law" or the more recent "radiation law" both calibrated using the real data available (*7*). These approaches assume more frequent travel to nearby or closer subpopulations and less frequent travel to distant locations. In Fig. 1 we show a representation of the geographical resolution of the model and the mobility network for the contiguous United States (US).

Initial conditions are set specifying the number and location of individuals capable of trans-

Parameters	Range	Ref.
Latent period (mean)	[4, 7] days	(*8*)
Infectious period (mean)	[2, 4] days	(*9*)
Days until recovery	[10, 14] days	(*9, 10*)
Generation time	[6, 8] days	
Reproductive number	[2.0, 3.1] in steps of 0.01	
Starting date	[2019-11-15, 2019-12-01]	(*11–14*)

Table 1: Summary of parameter rangs explored in the implementation of the model in GLEAM. Baseline parameters are reported in the main text.

mitting the infection. GLEAM is then able to track over time the proportion of the population in each disease compartment for all subpopulations. At the start of each simulated day, travelers move to their destinations via the flight network. The probability of air travel changes from day to day, varies by age group, and can consider the effects of location specific airline traffic reductions. Short-range mobility (i.e. commuting) varies between workdays and weekends, by age group, and by disease status. Each full day is simulated using 12 distinct time steps, and this process is repeated for every simulated day. Individuals and their traveling patterns are tracked as shown in the flow diagram for the GLEAM algorithm (Fig. 2).

The combined population structure and mobility network create a synthetic world that is used to simulate the unfolding dynamics of the epidemic. The infection dynamics occur within each subpopulation. We adopt a classic $SLIR$ model in which individuals can be classified into four compartments: susceptible, latent, infectious, or removed. Susceptible individuals become latent through interactions with infectious individuals. During both the latent and infectious stages we assume that individuals are able to travel. Following the infectious period, individuals then progress into the removed compartment where they are no longer able to infect others, meaning they have either recovered, been hospitalized, isolated or have died. Individuals transition between compartments using stochastic binomial chain processes assuming parameter values from available literature that define the natural history of disease. In Table. 1 we report the parameter estimates used in the model.

Interventions Timeline

In order to realistically depict the evolution of the epidemic, a comprehensive set of policy interventions is applied to modify disease transmissibility and population mobility. On January 15, partial international travel reductions (from 10% to 40%) are applied for individuals traveling to/from China. On January 23 and 25, flight and commuting reductions are applied respectively to the Wuhan subpopulation to enforce its state of quarantine with respect to the rest of China.

In addition, on January 25, commuting reductions are applied also to all other subpopulations in mainland China. To do so, we collected daily travel data starting January 1, 2020 until

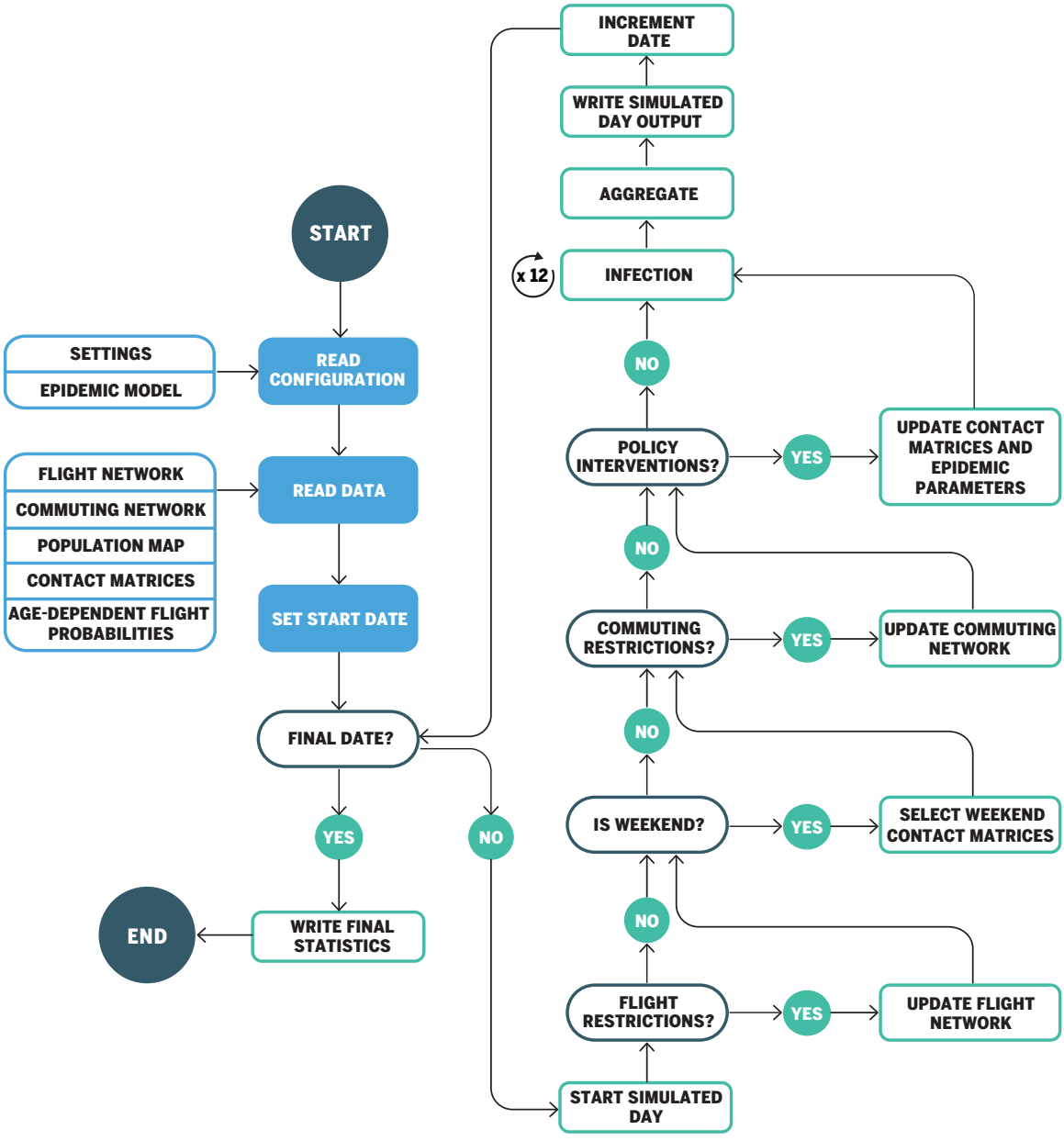

Figure 2: Flow diagram of GLEAM's algorithm

February 25, 2020 from the Baidu Qianxi platform (*15*), which provides three mobility indices (inflow index, outflow index, and intra-city index). The indices are proxies for the number of travelers moving in, out of, and inside a city, respectively. We extracted the mobility outflow index of 27 provinces and 4 municipalities for the current year 2020 and the previous year (with the same lunar date), and then mapped all provinces and municipalities to the metapopulation structure of the model to estimate the travel flow changes during the epidemic where the travel reduction can be estimated as $1 - \frac{I_{cur}}{I_{pre}}$, where I_{cur} and I_{pre} are the mobility outflow index of current year 2020 and previous year on the same lunar date, respectively.

On February 1, due to the increasing amount of restrictions implemented by various countries and airlines (*16–21*), stronger travel reductions are applied between mainland China and the rest of the world. This restriction corresponds to the executive order released on January 31, limiting travel between the US and mainland China (*22*). The reductions are country/territory specific and they range from a 91% reduction to a 100% reduction in the case of Hong Kong and Singapore. In addition, a 50% commuting reduction is applied whenever a subpopulation observes more than 10 cumulative infections per 100,000 individuals in the population. On March 4, we start to include travel restrictions related to travel advisories and the change of testing policy not based on travel history. Lastly, 50% and 80% average transmissibility reductions are assumed in mainland China starting, respectively, on January 25 and February 5 to capture the effects of local social distancing policies.

Global Model Calibration

The model described is stochastic and outputs an ensemble of possible epidemic outcomes for each set of initial conditions. We seed the epidemic in Wuhan, China assuming a starting date between 11/15/2019 and 12/1/2019 with 40 initial infections. Given the doubling time of the epidemic this might corresponds to the epidemic starting from mid October to late November, 2019. We simulate epidemic scenarios sampling reproductive numbers (R_0) from a uniform prior in the range 2.0 to 3.1. We calculate the distribution $P(D)$, where D is the evidence, $P(D|R_0)$, the likelihood of R_0, and $P(R_0|D)$ the posterior distribution of R_0. The evidence, D, is the growth rate of internationally imported cases during the exponential growth period of the epidemic in China. From the simulated epidemics, we select the scenarios where the cumulative number of detected, imported cases by January 23, 2020 matches the reported cases observed. The detailed list of importation events used is provided in Table S1 of the supplementary materials of Ref. (*23*). The $P(R_0 = x|D)$ is the fraction of simulations where both $R_0 = x$ and the evidence constraint on the detected importations holds over the total number of simulations where $R_0 = x$. Using the ABC calibration and the age-stratified contact matrices for the US, the obtained posterior distribution for the basic reproductive number R_0 in the US has a median 2.9 [95% CI 2.6-3.1] (3). The median values for state level reproductive numbers range from $2.8 - 3.0$, with doubling times T_d in the range of $3.1 - 3.7$ days.

Sensitivity Analysis

In the main text we report the results of the model calibrated only on importation data from mainland China before the implementation of any travel restrictions. Here we show the results

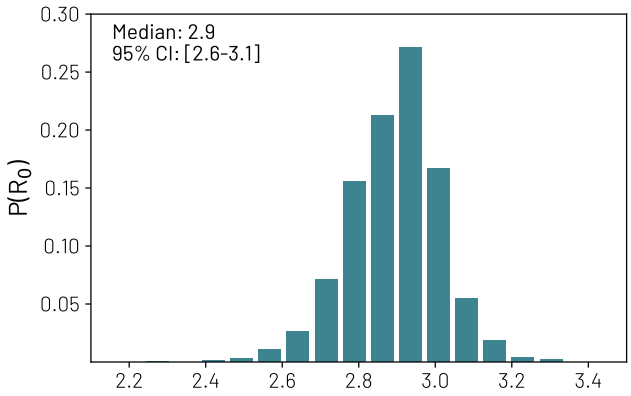

Figure 3: The posterior distribution of the reproductive number for the US. The average value and the 95% reference range are shown.

using the same compartmental model and parameters, however we calibrate the model using both the international importation events as well as an additional constraint on the weekly incidence of deaths reported in the US from March 21 to March 28, 2020. The deaths recorded in this period are most likely from infections that occurred before local interventions were implemented. The additional calibration further selects the simulation results that more closely match the evolution of the outbreak in the US. The average reproductive number R_0 for the individual states ranges $2.9 - 3.0$ with an average doubling time of 3.3 [90%CI 3.1-3.7] days. Both calibrations produce similar results and do not alter the conclusions of the main analysis. In Fig. 4 we show the distribution of the first dates when each continental state observed at least 10 locally generated transmission events per day. Model differences between the median date estimates between this calibration and the one in the main text are at most 5 days.

Epidemic surveillance data

The data of officially confirmed cases and deaths for each state is taken from John Hopkins University Coronavirus resource centre (*24*). The data used in the analysis has been downloaded on May 23 and contains data up to that date.

Importation networks

The importation networks are obtained as follows. As a first step, we track the importations by air transportation (considering both individuals in the latent and infectious compartments) in any census areas of the US in all the runs selected using the ABC calibration. We then compute the day, in each run, in which the number of daily transitions from S to L is at least 10 in each state. In order words, we evaluate the date, in each run, when the state experienced the first local outbreak. We then track, in each run, the arrivals of latent and infectious individuals before or at the time of the local outbreak. In doing so, we aggregate the observations from census areas to states. Then in each build, for each state, we construct a directed and weighted network in which importation sources link the target states. The width of the link is the number of infections imported. By considering the observations in all the runs selected during the calibration phase,

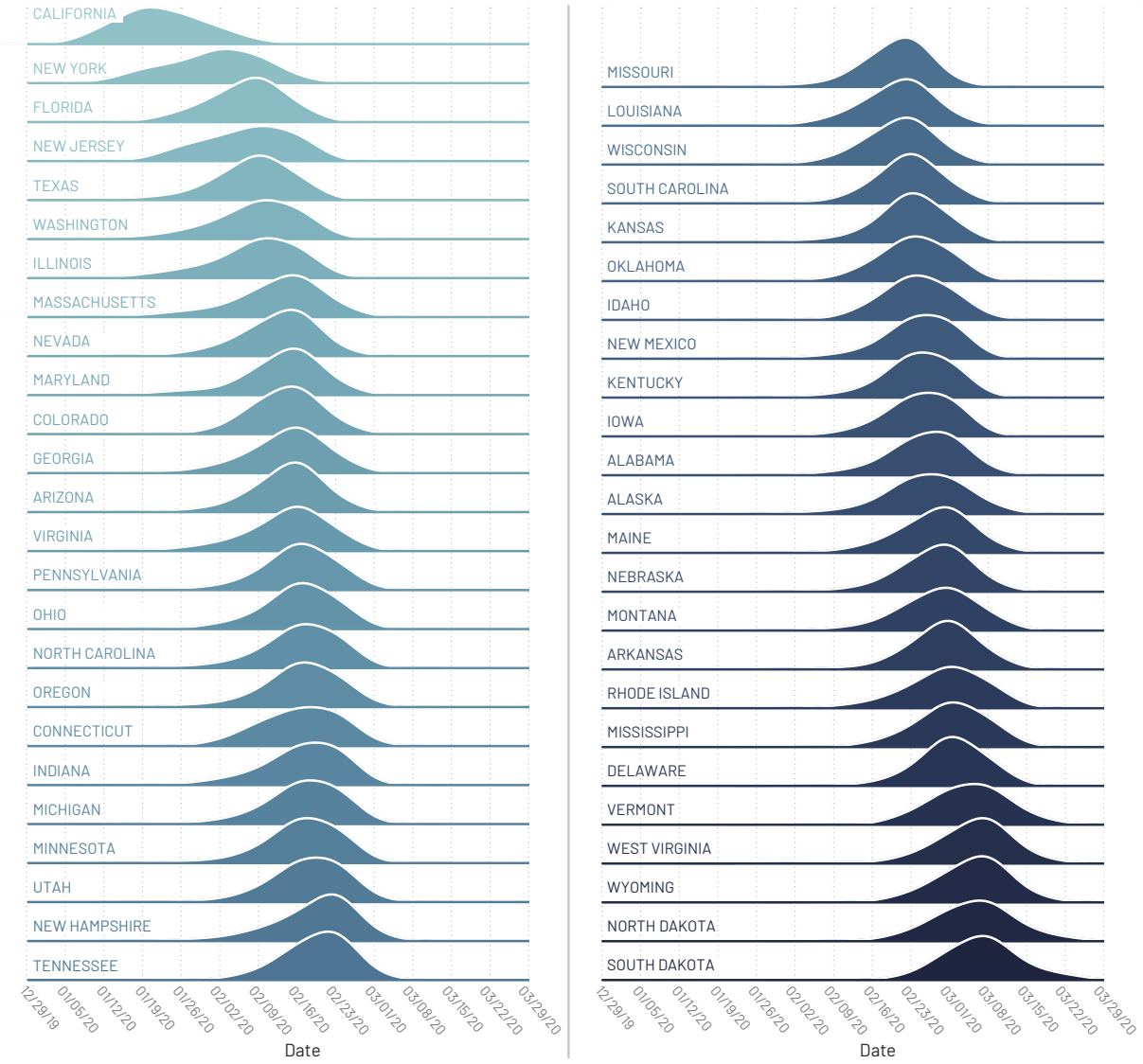

Figure 4: Timing of the onset of local transmission. Distributions of the probability that each state had reached at least 10 locally generated COVID-19 transmission events per day by any given week since December 1, 2019.

we build the chord diagram shown in the manuscript. To make them comparable, as the number of importations is a function of the source-target pair, we normalized the incoming flows for each target such that the sum of all incoming links is one. Furthermore, to help the readability of the plots, we aggregated sources considering macro areas such as Europe and Asia. We keep US (to isolate the national importations) and mainland China (as the epicenter of the pandemic) separate. All the other sources are grouped together and labeled "Others". More specifically, source countries of importations in the USA are grouped as:

- **Asia**: Afghanistan, Armenia, Azerbaijan, Bahrain, Bangladesh, Brunei, Cambodia, Cyprus, India, Indonesia, Iran, Iraq, Israel, Japan, Jordan, Kazakhstan, Korea, Kuwait, Kyrgyzstan, Lao PDR, Lebanon, Malaysia, Maldives, Mongolia, Myanmar, Nepal, Oman, Pakistan, Philippines, Qatar, Saudi Arabia, Singapore, Sri Lanka, Taiwan, Tajikistan, Thailand, Turkey, United Arab Emirates, Uzbekistan, Vietnam

- **China**: mainland China

- **Europe**: Albania, Austria, Belarus, Belgium, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Jersey, Kosovo, Latvia, Lithuania, Luxembourg, Macedonia, Malta, Moldova, Montenegro, Netherlands, Norway, Poland, Portugal, Romania, Russian Federation, Serbia, Slovak Republic, Slovenia, Spain, Sweden, Switzerland, Ukraine, United Kingdom

- **Others**: Algeria, Angola, Antigua and Barbuda, Argentina, Aruba, Australia, Bahamas, Barbados, Belize, Bermuda, Botswana, Brazil, British Virgin Islands, Canada, Cape Verde, Caribbean Netherlands, Cayman Islands, Chile, Colombia, Congo, Cook Islands, Costa Rica, Cuba, Curaçao, Côte dIvoire, Djibouti, Dominica, Dominican Republic, Ecuador, Egypt, El Salvador, Equatorial Guinea, Ethiopia, Fiji, French Guiana, French Polynesia, Gabon, Gambia, Ghana, Greenland, Grenada, Guadeloupe, Guatemala, Guinea, Guyana, Haiti, Honduras, Jamaica, Kenya, Liberia, Madagascar, Mali, Marshall Islands, Martinique, Mauritius, Mexico, Morocco, Namibia, New Zealand, Nicaragua, Nigeria, Palau, Panama, Paraguay, Peru, Rwanda, Samoa, Senegal, Seychelles, South Africa, St-Barthélemy, St. Kitts and Nevis, St. Lucia, St. Maarten, St. Vincent and Grenadines, Sudan, Suriname, Tonga, Trinidad and Tobago, Tunisia, Turks and Caicos Islands, Uganda, Uruguay, Venezuela, Zambia, Zanzibar, Zimbabwe

- **USA**: all the US states plus the US territories (American Samoa, Guam, Northern Mariana Islands, Puerto Rico, U.S. Virgin Islands)

Following this procedure, in Table 2 we report the seeding share for the first 25 states that experienced a local outbreak (same states as in the left column of Fig. 2 in the manuscript).

In Table 3 we report the seeding shares for the remaining states that experienced a local outbreak (same states as in the right column of Fig. 2 in the manuscript).

State	USA	China	Asia	Europe	Others
California	0.19	0.42	0.25	0.04	0.10
New York	0.32	0.15	0.26	0.16	0.11
Florida	0.57	0.02	0.12	0.14	0.16
Washington	0.58	0.03	0.23	0.05	0.10
New Jersey	0.37	0.07	0.26	0.18	0.13
Texas	0.59	0.02	0.21	0.07	0.11
Illinois	0.55	0.04	0.21	0.09	0.11
Massachusetts	0.52	0.03	0.18	0.15	0.12
Maryland	0.53	0.02	0.23	0.13	0.10
Nevada	0.61	0.01	0.16	0.07	0.15
Georgia	0.66	0.01	0.16	0.08	0.09
Virginia	0.58	0.01	0.20	0.11	0.09
Arizona	0.71	$<$0.01	0.10	0.04	0.14
Colorado	0.72	$<$0.01	0.09	0.06	0.12
Ohio	0.69	0.01	0.15	0.06	0.09
Pennsylvania	0.69	0.01	0.10	0.09	0.11
North Carolina	0.69	0.01	0.14	0.08	0.09
Michigan	0.64	0.01	0.18	0.08	0.09
Minnesota	0.68	$<$0.01	0.13	0.07	0.12
Connecticut	0.49	0.01	0.21	0.16	0.13
Indiana	0.67	0.01	0.15	0.07	0.10
Oregon	0.70	0.01	0.17	0.03	0.09
Utah	0.77	$<$0.01	0.10	0.04	0.09
New Hampshire	0.55	0.01	0.16	0.15	0.12
Tennessee	0.73	$<$0.01	0.11	0.06	0.10

Table 2: Importation of seeding events. Sources are listed from the second column on. Targets are the states listed in the first column. This is the list of the first 25 states that experienced an outbreak. Numbers are rounded to the second digit.

State	USA	China	Asia	Europe	Others
Missouri	0.77	<0.01	0.09	0.05	0.09
Louisiana	0.76	<0.01	0.08	0.06	0.10
Wisconsin	0.83	<0.01	0.05	0.03	0.09
South Carolina	0.77	<0.01	0.09	0.07	0.07
Kansas	0.81	<0.01	0.07	0.05	0.07
Oklahoma	0.78	<0.01	0.09	0.05	0.08
Kentucky	0.76	<0.01	0.11	0.05	0.07
Idaho	0.83	<0.01	0.09	0.02	0.06
New Mexico	0.86	<0.01	0.05	0.03	0.05
Iowa	0.82	<0.01	0.06	0.04	0.09
Alabama	0.79	<0.01	0.11	0.04	0.06
Alaska	0.74	<0.01	0.18	0.02	0.06
Maine	0.82	<0.01	0.06	0.06	0.06
Nebraska	0.85	<0.01	0.05	0.03	0.07
Montana	0.92	<0.01	0.01	0.02	0.04
Arkansas	0.86	<0.01	0.05	0.04	0.05
Rhode Island	0.90	<0.01	0.02	0.05	0.03
Delaware	0.79	<0.01	0.05	0.08	0.08
Mississippi	0.86	<0.01	0.06	0.04	0.04
Vermont	0.93	<0.01	0.03	0.02	0.02
North Dakota	0.95	0.	0.01	0.01	0.03
Wyoming	0.91	<0.01	0.02	0.02	0.05
West Virginia	0.91	<0.01	0.03	0.03	0.03
South Dakota	0.94	0.	0.01	0.02	0.04

Table 3: Importation of seeding events. Sources are listed from the second column on. Targets are the states listed in the first column. This is the list of the last 24 states that experienced an outbreak. Numbers are rounded to the second digit.

In Table 4 we report the the share of introduction of SARS-Cov-2 infections for all the first 25 states that experienced a local outbreak (same states as in the left column of Fig. 2 in the manuscript), considering, however, all the infections imported up to March 1, 2020. In other words, we don't constrain on the time of the start of the local outbreak (i.e. first daily 10 local transmissions), but we consider a general cutoff date for all states. The flows are radically different, especially for the first states, where the critical role of China before the travel restrictions of January 23, is replaced by a much larger internal flow. In fact, for all states the share of domestic importations is above 50%.

In Table 5 we report the share of introduction of SARS-Cov-2 infections through March 1, 2020, for the remaining states that experienced a local outbreak (same states in the right column of Fig. 2 in the manuscript).

In Figure 5 we provide a visual representation, via a chord plot, of the data shown in Tables 4 and 5.

Correlation analysis

As mentioned and shown in the main text, during the early phases of the spreading, mobility plays a crucial role. In order to highlight this aspect, here we report the full correlation analysis between the real data and the mobility indicators. In particular, we compute the order in which states reached 100 cases/deaths in the real surveillance data and compare it with the order of states according to their air traffic (considering both national and international travels). Note how the correlation plot in the main text considered as a mobility indicator the sum of the two types of traffic. In Fig. 6 we show the result reporting also the value of the Kendall's tau. Overall the correlations, especially for the number of cases, are comparable with those shown in the main text.

The states that were the first to experience the outbreak, besides being hubs in the air transportation network, are also very populous. It is then natural to wonder how rankings based on population compare with respect to those based on air traffic. In Figure 7 we show the comparison. In particular, we order states according to their population (left column) and density (right column) and to the epidemic indicators from surveillance (cases and deaths). Not surprisingly, we find high correlation levels across the board. However, it is interesting to note how those reported considering air travels are comparable (0.67 when considering both national and international traffic). The correlation for the number of deaths (bottom row) is lower with respect to the number of cases (top row). Furthermore, it is interesting to notice how the correlations are even smaller when considering the population density (right column). In fact, they go down to 0.38 when comparing them with the order in which a state surpassed the threshold of 100 deaths.

In Figure 8 we repeat the same analysis considering the model's projections. The correlations are comparable to the previous. Also in the model, population density is less correlated. It is important to observe that air travel traffic, population and population densities are not independent indicators. Figure 9 highlights this observation. Particularly high is the correlation between air traffic and population. Also, population and population densities are well correlated

State	USA	China	Asia	Europe	Others
California	0.64	<0.01	0.18	0.05	0.12
New York	0.52	<0.01	0.17	0.18	0.13
Florida	0.71	<0.01	0.05	0.09	0.15
Washington	0.79	<0.01	0.11	0.03	0.07
New Jersey	0.53	<0.01	0.16	0.17	0.14
Texas	0.79	<0.01	0.09	0.04	0.08
Illinois	0.73	<0.01	0.10	0.07	0.10
Massachusetts	0.68	<0.01	0.09	0.12	0.11
Maryland	0.70	<0.01	0.12	0.09	0.09
Nevada	0.76	<0.01	0.07	0.04	0.12
Georgia	0.80	<0.01	0.07	0.05	0.07
Virginia	0.74	<0.01	0.10	0.08	0.08
Arizona	0.82	<0.01	0.04	0.02	0.12
Colorado	0.85	<0.01	0.04	0.03	0.08
Ohio	0.82	<0.01	0.07	0.04	0.07
Pennsylvania	0.81	<0.01	0.05	0.06	0.09
North Carolina	0.82	<0.01	0.06	0.05	0.07
Michigan	0.77	<0.01	0.10	0.05	0.07
Minnesota	0.80	<0.01	0.06	0.04	0.10
Connecticut	0.63	<0.01	0.12	0.13	0.12
Indiana	0.79	<0.01	0.08	0.05	0.08
Oregon	0.84	<0.01	0.08	0.02	0.06
Utah	0.87	<0.01	0.05	0.02	0.06
New Hampshire	0.67	<0.01	0.10	0.12	0.11
Tennessee	0.82	<0.01	0.06	0.04	0.08

Table 4: Introduction of SARS-Cov-2 infections thorugh March 1. Sources are listed from the second column on. Targets are the states listed in the first column. This is the list of the first 25 states that experienced an outbreak. Numbers are rounded to the second digit.

State	USA	China	Asia	Europe	Others
Missouri	0.85	<0.01	0.04	0.03	0.07
Louisiana	0.84	<0.01	0.04	0.04	0.08
Wisconsin	0.88	<0.01	0.02	0.02	0.08
South Carolina	0.84	<0.01	0.05	0.05	0.06
Kansas	0.87	<0.01	0.04	0.03	0.06
Oklahoma	0.85	<0.01	0.05	0.03	0.07
Kentucky	0.82	<0.01	0.08	0.04	0.07
Idaho	0.89	<0.01	0.05	0.01	0.05
New Mexico	0.90	<0.01	0.03	0.02	0.04
Iowa	0.85	<0.01	0.04	0.03	0.09
Alabama	0.82	<0.01	0.09	0.04	0.05
Alaska	0.84	<0.01	0.11	0.01	0.04
Maine	0.84	<0.01	0.05	0.05	0.07
Nebraska	0.85	<0.01	0.04	0.02	0.08
Montana	0.93	<0.01	0.01	0.01	0.05
Arkansas	0.87	<0.01	0.04	0.03	0.05
Rhode Island	0.86	<0.01	0.02	0.08	0.04
Delaware	0.78	<0.01	0.04	0.07	0.10
Mississippi	0.84	<0.01	0.06	0.04	0.05
Vermont	0.90	<0.01	0.04	0.02	0.04
North Dakota	0.92	0.	0.01	0.01	0.06
Wyoming	0.87	<0.01	0.03	0.02	0.09
West Virginia	0.87	<0.01	0.04	0.03	0.05
South Dakota	0.90	0.	0.01	0.02	0.07

Table 5: Importation of infection flows up to March 1. Sources are listed from the second column on. Targets are the states listed in the first column. This is the list of the last 24 states that experienced an outbreak. Numbers are rounded to the second digit.

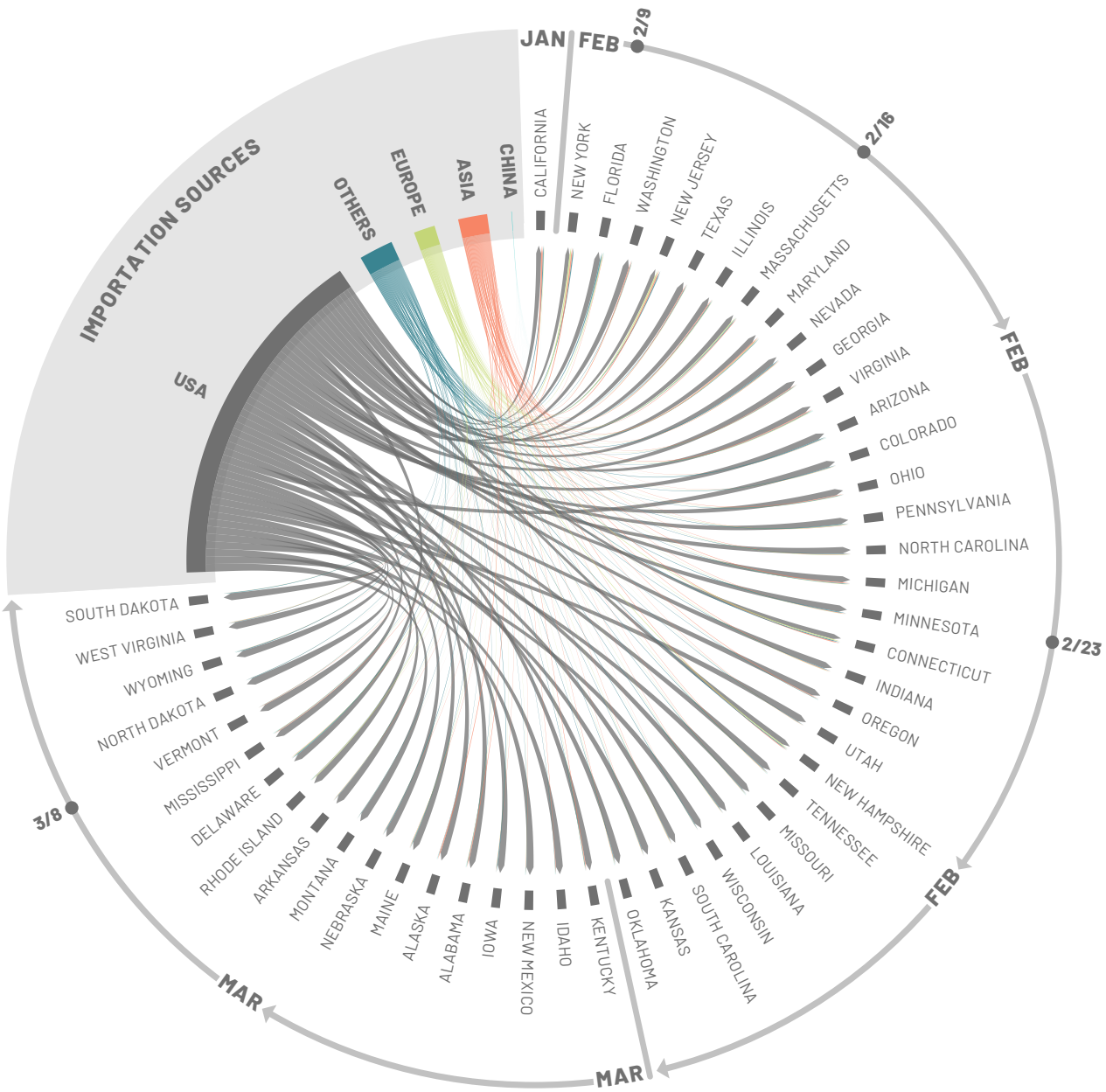

Figure 5: Share of importations of infections in all continental states from US, China, Europe, Asia and all other countries up to March 1, 2020. US states are ordered, clockwise, according to the start of the local outbreak.

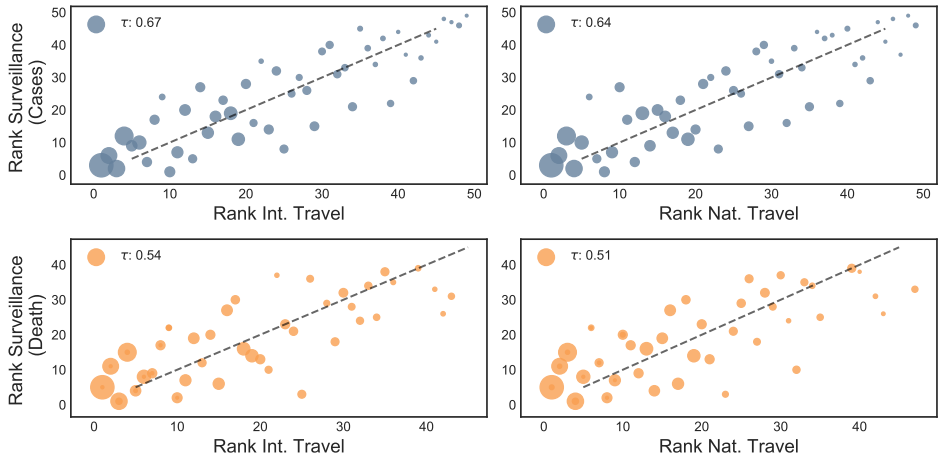

Figure 6: Correlation between the order in which states reached 100 confirmed cases (top row) or deaths (bottom row) and their International (left) or National (right) air traffic. The size of each state is assigned proportional to the population size

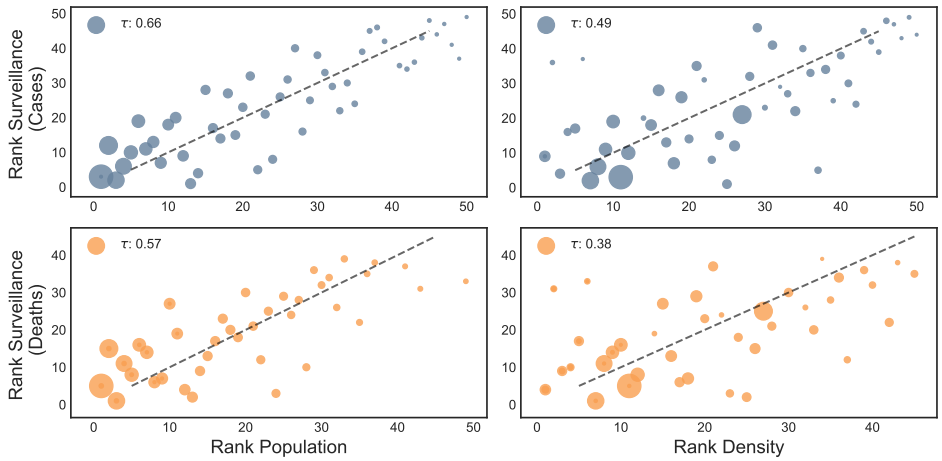

Figure 7: Correlation between the order in which states reached 100 confirmed cases (top row) or deaths (bottom row) and their population (left) or population density (right). The size of each state is assigned proportional to the population size

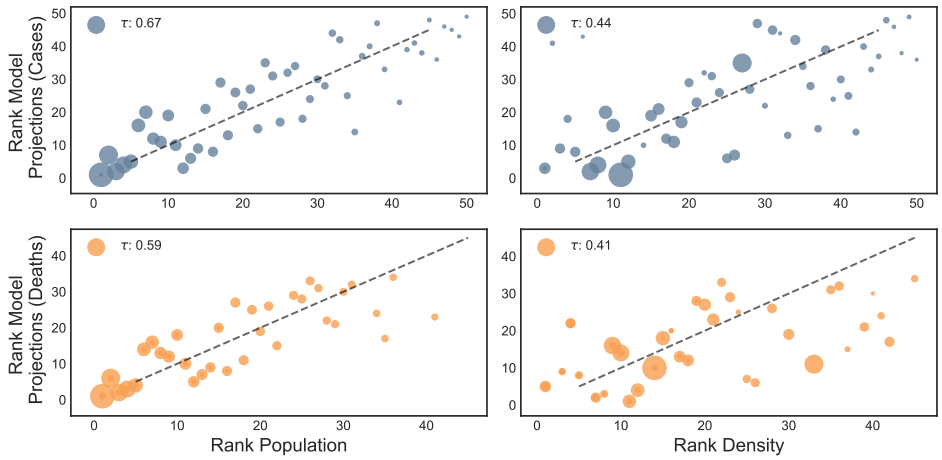

Figure 8: Correlation between the order in which states reached 100 confirmed cases (top row) or deaths (bottom row) according to the model and their population (left) or population density (right). The size of each state is assigned proportional to the population size

while air traffic and density are not. This is due, in part, to the many states that, due to their location, see lots of traffic but are not very dense.

The correlations between rankings reported above have been computed by using the Kendall's tau (*25*) as implemented by the *scipy.stats* library (*26*). The metric is designed to compare the rankings obtained ordering items, states in our case, according to pairs of different quantities. The Kendall's tau is defined only in the case that the ranks have the same size. In case the two ranks have different size (i.e. some states did not yet go above a given threshold) the metric is applied to the common subset of the two.

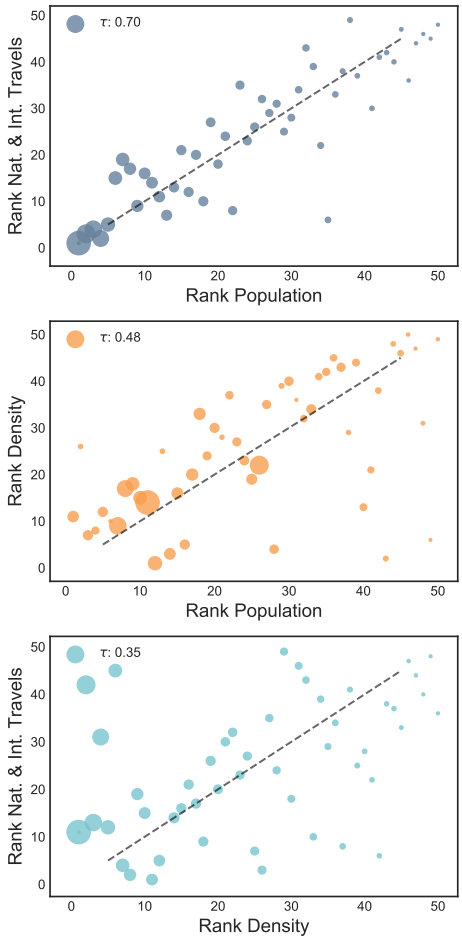

Figure 9: Correlation between travel, population and population density for the continental US states.

References

1. D. Balcan, B. Gonçalves, H. Hu, J.J. Ramasco, V. Colizza, A. Vespignani, Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model. *Journal of Computational Science* **1**, 132-145 (2010).

2. D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J.J. Ramasco, A. Vespignani, Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences*. **106**, 21484-21489 (2009).

3. Socioeconomic Data and Applications Center (SEDAC), Columbia University. http://sedac.ciesin.columbia.edu/gpw.

4. D. Mistry, M. Litvinova, M. Chinazzi, L. Fumanelli, M. F. Gomes, S. A. Haque, Q.-H. Liu, K. Mu, X. Xiong, M. E. Halloran, I.M. Longini Jr., S. Merler, M. Ajelli, A. Vespignani. Inferring high-resolution human mixing patterns for disease modeling. *arXiv* [Preprint]. 25 February 2020. https://arxiv.org/abs/2003.01214

5. International Air Transportation Association. https://www.iata.org/.

6. Official Aviation Guide. https://www.oag.com/.

7. F. Simini, M. C. González, A. Maritan, A.-L. Barabási, *Nature* **484**, 96-100 (2012).

8. J. A. Backer, D. Klinkenberg, J. Wallinga, Incubation period of 2019 novel coronavirus (2019-nCoV) infections among travellers from Wuhan, China, 20–28 January 2020.*Eurosurveillance* **25**(5), 2000062 (2020). https://doi.org/10.2807/1560-7917.ES.2020.25.5.2000062

9. S. M. Kissler, C. Tedijanto, E. Goldstein, Y. H. Grad, M. Lipsitch, Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period *Science* **368**, 860 (2020).

10. R. Verity, L. C. Okell, I. Dorigatti, P. Winskill, C. Whittaker, N. Imai, G. Cuomo-Dannenburg, H. Thompson, P. G. T. Walker, H. Fu, A. Dighe, J. T. Griffin, M. Baguelin, S. Bhatia, A. Boonyasiri, A. Cori, Z. Cucunubá, R. FitzJohn, K. Gaythorpe, W. Green, A. Hamlet, W. Hinsley, D. Laydon, G. Nedjati-Gilani, S. Riley, S. van Elsland, E. Volz, H. Wang, Y. Wang, X. Xi, C. A. Donnelly, A. C. Ghani, N. M. Ferguson. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet Infectious Diseases* (2020). https://doi.org/10.1016/S1473-3099(20)30243-7

11. A. Rambaut, "Preliminary phylogenetic analysis of 11 nCoV2019 genomes, 2020-01-19" (2020); http://virological.org/t/preliminary-phylogenetic-analysis-of-11-ncov2019-genomes-2020-01-19/329.

12. N. Imai, A. Cori, I. Dorigatti, M. Baguelin, C. A. Donnelly, S. Riley, N. M. Ferguson, "Report 3: Transmissibility of 2019-nCoV" (Imperial College London, 2020) www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/report-3-transmissibility-of-covid-19/.

13. K. Anderson, “Estimates of the clock and TMRCA for 2019-nCoV based on 27 genomes” (2020); http://virological.org/t/clock-and-tmrca-based-on-27-genomes/347

14. T. Bedford, R. Neher, J. Hadfield, E.Hodcroft, M. Ilcisin, N. Müller, “Genomic analysis of nCoV spread. Situation report 2020-01-23” (2020); https://nextstrain.org/narratives/ncov/sit-rep/2020-01-23

15. Baidu Qianxi, http://qianxi.baidu.com/ (2020).

16. New York Times, “North Korea Bans Foreign Tourists Over Coronavirus, Tour Operator Says”, https://www.nytimes.com/2020/01/21/world/asia/coronavirus-china-north-korea-tourism-ban.html (2020).

17. CNA, “Scoot cancels flights to China’s Wuhan over virus outbreak”, https://www.channelnewsasia.com/news/singapore/wuhan-virus-scoot-cancels-flights-mtr-train-12309076 (2020).

18. Toui tre News, “Vietnam aviation authority ceases all flights to and from coronavirus-stricken Wuhan”, https://tuoitrenews.vn/news/business/20200124/vietnam-aviation-authority-ceases-all-flights-to-and-from-coronavirusstricken-wuhan/52707.html (2020).

19. Reuters, “Russia ramps up controls, shuts China border crossings over virus fears”, https://www.reuters.com/article/us-china-health-russia-border/russian-regions-in-far-east-close-border-with-china-amid-coronavirus-fears-tass-idUSKBN1ZR0TU (2020).

20. Center for Disease Control, “Novel Coronavirus in China”, https://wwwnc.cdc.gov/travel/notices/warning/novel-coronavirus-china (2020).

21. The Australian, “Travelers from China to be denied entry to Australia”, https://www.theaustralian.com.au/subscribe/news/1/?sourceCode=TAWEB_WRE170_a&dest=https%3A%2F%2Fwww.theaustralian.com.au%2Fnation%2Ftravellers-from-china-to-be-denied-entry-into-australia%2Fnews-story%2F7b7619d44af78dd7395a934e22b52997&memtype=anonymous&mode=premium (2020).

22. Proclamation on Suspension of Entry as Immigrants and Nonimmigrants of Persons who Pose a Risk of Transmitting 2019 Novel Coronavirus, https://www.whitehouse.gov/presidential-actions/proclamation-suspension-entry-immigrants-nonimmigrants-persons-pose-risk-transmitting-2019-novel-coronavirus/ (2020).

23. M. Chinazzi, J.T. Davis, M. Ajelli, C. Gioannini, M. Litvinova, S. Merler, A.P. y Piontti, L. Rossi, K. Sun, C. Viboud, X. Xiong, H. Yu, E.M. Halloran, I.M. Longini, A. Vespignani, The effect of travel restrictions on the spread of the 2019 novel coronavirus (COVID-19) outbreak. *Science* **368**, 395–400 (2020).

24. John Hopkins University Coronavirus resource centre https://coronavirus.jhu.edu/.

25. M. G. Kendall, A new measure of rank correlation. *Biometrika* **30**, 81 (1938).

26. Scipy.org: Kendall Tau. https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.kendalltau.html.