Assessment of the similarities and differences of the COVID-19 second surges in Europe and the Northeast United States

Jose M. G. Vilar[1,2,*] and Leonor Saiz[3,*]

[1] Biofisika Institute (CSIC, UPV/EHU), University of the Basque Country (UPV/EHU), P.O. Box 644, 48080 Bilbao, Spain

[2] IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Spain

[3] Department of Biomedical Engineering, University of California, 451 E. Health Sciences Drive, Davis, CA 95616, USA

[*]Correspondence to: j.vilar@ikerbasque.org or lsaiz@ucdavis.edu

ABSTRACT

We estimated the aggregate prevalence over time for Europe and the Northeast United States to characterize the COVID-19 second surge in these regions. We find a starting date as early as July 3 for Europe and August 19 for the Northeast; subsequent infectious populations that, as of January 07, 2021, have always increased or remained stagnant; and the resurgences being the collective effect of each overall region with no location dominating the dynamics by itself.

INTRODUCTION

Europe and the Northeast United States (US) are progressing through a second surge of the COVID-19 outbreak after experiencing a similar initial outburst and subsequent regression. None of these resurgences was widely expected nor anticipated [1, 2]. Both regions display high mobility among their locations and broad independence among locations to enact measures to mitigate the propagation of the outbreak. In the case of Europe, Schengen Area countries allow for unrestricted border crossings among them. Mobility restrictions,

lockdowns, and other nonpharmaceutical interventions were able to achieve a major regression of the outbreaks, but the gradual lifting of restrictions has resulted in a resurgence across locations in these two regions [3, 4]. The characterization of the similarities and differences of the outbreak progression in these two areas is needed to provide insights into the effectiveness of the actions taken, to ascertain the extent their results can be extrapolated from one region to another, and to informedly mitigate the current and potentially forthcoming resurgences.

Surveillance at the initial stages of a resurgence faces high uncertainty because of the low prevalence values at which it happens [4-7]. Primarily, direct characterization through randomized testing would need large population studies to provide significant results and using infection case data is dependent on varying testing rates [4, 6]. More robust approaches based on death counts are also affected by the extremely small number of random events on which they rely for inference [4, 7]. This uncertainty in assessing the state of the outbreak for a potential resurgence is a source of delays in the decision making and intervention implementation processes.

METHODS

Aggregate values provide a potential avenue to increase the reliability of the estimates for low prevalence values

To assess the properties of the resurgences with increased confidence, we computed the aggregate values of the infectious populations and the corresponding time-varying reproduction numbers for Europe's Schengen Area and the Northeast US from the individual values of the locations of each region [7]. We considered overall region values and overall region values excluding one location. Exclusion of one location provides an avenue to reliably infer the effects of the location in the overall region.

Reproduction numbers were computed from growth rates considering a gamma-distributed generation interval with a mean of 6.5 days and a standard deviation of 4.2 days (Appendix) [8]. Aggregate values were computed from infectious-population-weighted individual location values (Appendix). The individual location data was downloaded on

January 30, 2021 from https://github.com/Covid19Dynamics/trajectories. The data considers explicitly the age-specific infection fatality rates from Verity et al. [9], which are consistently similar among distinct locations [10, 11], to infer the local infectious population from reported death counts [7].

RESULTS

The second surges started in early-mid summer with the Northeast US trailing Europe

The initial progression of the overall infectious populations for both regions consisted of exponential growth followed by exponential decay (Figures 1A and 1B). Subsequently, there was a sharp transition to fast exponential growth in Europe on July 3 (95% Confidence Interval (CI): July 1–July 6), 2020, from an estimated infectious population of 3.0×10^4 (95 % CI: 2.4×10^4 – 3.5×10^4) individuals and a stagnant overall infectious population in the Northeast US, which started to grow slowly but with increasing speed on August 19 (95% CI: August 16 – August 23), 2020, from an estimated infectious population of 3.3×10^4 (95 % CI: 2.8×10^4 – 3.8×10^4) individuals.

The resurgence has been more abrupt and intense in Europe than in the Northeast US

Concomitantly, the time-varying reproduction numbers crossed above one on the resurgence dates less abruptly in the Northeast US than in Europe (Figure 1C), reaching maximum values of 1.50 (95% CI: 1.48–1.51) in Europe and 1.30 (95% CI: 1.27–1.34) in the Northeast US. The sharp resurgence to exponential growth in Europe is coincidental with lifting major nonpharmaceutical interventions that curved the outbreak [12], including the coordinated end of travel bans in Schengen Area's countries on July 1, 2020 [13].

There is no evidence of an incipient regression in any of the two surges as of January 07, 2021

No substantial decreases in the overall infectious population, nor corresponding reproduction numbers below one, have been observed for any of the two regions after the starting dates of the second surges (Figure 1). As of January 07, 2021, the estimated

infectious population has just stooped growth in the Northeast US (Figure 1B) and has remained in a stagnant state since early November in Europe (Figure 1A).

Aggregate values are highly reliable compared to location-specific data

The low prevalence at the specific-location level leads to broad confidence intervals for both the infectious population and the time-varying reproduction numbers, which makes ascertaining the local growth properties of the outbreak unreliable over prolonged periods of time (Supplementary Figure S1). The aggregated values for each region provide precise evidence of sustained growth of the outbreaks already over the summer, despite the uncertainty and variability present in each of the locations independently (Appendix, Figure 2, and Supplementary Figure S1).

The resurgences are the collective effect of each overall region

There is also robust evidence that the resurgence was not driven by a unique location since any aggregate value of the starting date for each region leaving one of their locations out is within the confidence limits of that of the overall region (Appendix, Figure 2, and Supplementary Figure S1). Therefore, the resurgence was the collective effect of each overall region.

DISCUSSION

These results highlight the need to implement policies and surveillance approaches that include data at a supra-location level when there is high mobility among locations. In this regard, the Northeast US, as a region, is closely trailing Europe in the second surge of the outbreak, but with a markedly smaller growth and evidence of slowing down earlier in the growth phase than Europe. Key differences in the actions taken included more gradual lifting and swifter progressive reimplementation of measures in the Northeast US than in Europe [12]. The progression over time of the aggregate prevalence of Europe's Schengen Area countries shows, with high confidence, that Europe's initial acting upon the second surge in mid-late October [1] took place well after a three-month-long period of sustained growth of the COVID-19 infectious population in the overall region, which has resulted in a second surge deadlier than the first one [14]. Such a high death toll has not taken place yet

in the Northeast US and, according to the observed trend in the progression of its estimated overall aggregate prevalence, it is potentially avertable with the social restrictions currently in place [12].

APPENDIX

The aggregate infectious population for a region is expressed as

$$n_I(t) = \sum_j n_{I,j}(t),$$

where $n_{I,j}(t)$ is the infectious population of the specific location with index j. The corresponding upper and lower confidence intervals are computed as

$$n_I^U(t) = n_I(t) + \sqrt{\sum_j \left(n_{I,j}^U(t) - n_{I,j}(t)\right)^2}$$

and

$$n_I^L(t) = n_I(t) - \sqrt{\sum_j \left(n_{I,j}(t) - n_{I,j}^L(t)\right)^2}$$

from the upper, $n_{I,j}^U(t)$, and lower, $n_{I,j}^L(t)$, confidence intervals for each location.

The overall time-dependent growth rate is given by

$$k_G(t) = \sum_j k_{G,j}(t) \frac{n_{I,j}(t)}{n_I(t)},$$

where $k_{G,j}(t)$ is the growth rate of the infectious population of the specific location with index j. This expression follows from

$$\frac{d}{dt} n_I(t) = \sum_j \frac{d}{dt} n_{I,j}(t) = \sum_j k_{G,j}(t)\, n_{I,j}(t) = k_G(t) n_I(t).$$

The corresponding upper and lower confidence intervals are computed as

$$k_G^U(t) = k_G(t) + \frac{1}{n_I(t)} \sqrt{\sum_j \left(k_{G,j}^U(t) - k_{G,j}(t)\right)^2 n_{I,j}(t)^2}$$

and

$$k_G^L(t) = k_G(t) - \frac{1}{n_I(t)} \sqrt{\sum_j \left(k_{G,j}(t) - k_{G,j}^L(t)\right)^2 n_{I,j}(t)^2}$$

from the upper, $k_{G,j}^U(t)$, and lower, $k_{G,j}^L(t)$, confidence intervals for each location.

The time-varying reproduction number R_t is obtained from the growth rate $k_G(t)$ through the Euler–Lotka equation,

$$R_t^{-1} = \int_0^\infty f_{GT}(\tau) e^{-k_G(t)\tau} d\tau,$$

where $f_{GT}(\tau)$ is the probability density function of the generation time. We consider the usual description of generation times through a gamma distribution

$$f_{GT}(\tau) = \frac{\beta^\alpha}{\Gamma(\alpha)} \tau^{\alpha-1} e^{-\beta\tau},$$

which leads to

$$R_t = \left(1 + \frac{k_G(t)}{\beta}\right)^\alpha$$

for $k_G(t) > -\beta$ and $R_t = 0$ for $k_G(t) \leq -\beta$. The values of the parameters are given by $\alpha = \tau_G^2/\sigma_G^2$ and $\beta = \tau_G/\sigma_G^2$, where τ_G and σ_G^2 are the mean and the variance of the generation time, respectively.

The starting date of the second surge is computed as the date the infectious population reached a minimum value after the maximum of the first surge. The lower bound of its confidence interval is computed as the last day before the minimum in which the upper bound of the confidence interval of the reproduction number is below 1. Analogously, the upper bound of the confidence interval is computed as the first day after reaching the minimum in which the lower bound of the confidence interval of the reproduction number is above 1.

The trajectories of the infectious populations, the growth rates, and the confidence intervals for each location were downloaded on January 30, 2021 from https://github.com/Covid19Dynamics/trajectories.

ACKNOWLEDGMENTS

J.M.G.V. acknowledges support from Ministerio de Ciencia e Innovación under grant PGC2018-101282-B-I00 (MCI/AEI/FEDER, UE). L.S. acknowledges support from the University of California, Davis.

REFERENCES

1. Looi, M.-K., *Covid-19: Is a second wave hitting Europe?* BMJ, 2020. **371**: p. m4113.
2. Kim, J., *Is a Second Wave Starting in N.Y.C.?* The New York Times 2020: p. www.nytimes.com/2020/11/10/nyregion/second-coronavirus-wave-.html.
3. Li, Y., et al., *The temporal association of introducing and lifting non-pharmaceutical interventions with the time-varying reproduction number (R) of SARS-CoV-2: a modelling study across 131 countries.* The Lancet Infectious Diseases 2020: p. DOI:10.1016/s1473-3099(20)30785-4.
4. Flaxman, S., et al., *Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe.* Nature, 2020. **584**(7820): p. 257-261.
5. Li, R., et al., *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV-2).* Science, 2020. **368**(6490): p. 489-493.
6. Vilar, J.M.G. and J.M. Rubi, *Determinants of population responses to environmental fluctuations.* Sci Rep, 2018. **8**(1): p. 887.
7. Vilar, J.M.G. and L. Saiz, *The evolving worldwide dynamic state of the COVID-19 outbreak.* medRxiv, 2020: p. 2020.11.26.20239434.
8. Bi, Q., et al., *Epidemiology and transmission of COVID-19 in 391 cases and 1286 of their close contacts in Shenzhen, China: a retrospective cohort study.* The Lancet Infectious Diseases, 2020. **20**(8): p. 911-919.
9. Verity, R., et al., *Estimates of the severity of coronavirus disease 2019: a model-based analysis.* Lancet Infect Dis, 2020. **20**(6): p. 669-677.

10. Levin, A.T., et al., *Assessing the age specificity of infection fatality rates for COVID-19: systematic review, meta-analysis, and public policy implications.* Eur J Epidemiol, 2020. **35**(12): p. 1123-1138.

11. O'Driscoll, M., et al., *Age-specific mortality and immunity patterns of SARS-CoV-2.* Nature, 2020: p. DOI:10.1038/s41586-020-2918-0.

12. Hale, T., et al. *Oxford COVID-19 Government Response Tracker, Blavatnik School of Government*. 2020. www.bsg.ox.ac.uk/covidtracker.

13. *Council of the European Union. Council Recommendation on the temporary restriction on non-essential travel into the EU and the possible lifting of such restriction*. 2020. https://data.consilium.europa.eu/doc/document/ST-9208-2020-INIT/en/pdf.

14. *JHU CSSE COVID-19 Data*. January 30, 2021; Available from: https://github.com/CSSEGISandData/COVID-19.

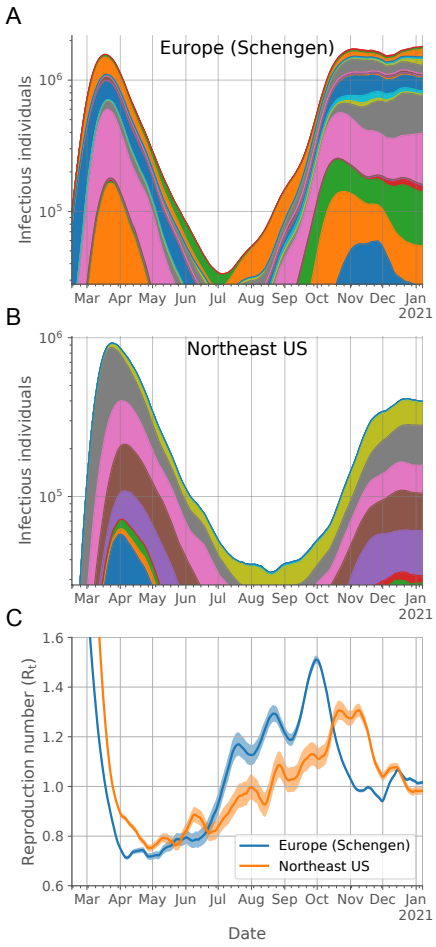

Figure 1. Temporal evolution of the COVID-19 outbreak in Europe and the Northeast US. Progression over time of the infectious population for countries in Europe's Schengen Area (A) and locations in the Northeast US (B) and of their corresponding reproduction numbers (C). Each colored section in the area plots represents the contribution of a country (A) or state (B) to the overall infectious populations. Countries and states are arranged in alphabetical order from bottom to top. The infectious populations are plotted on a logarithmic scale to highlight the triphasic behavior (growth-decay-growth) of the outbreak. The shaded regions in the reproduction number plots (C) represent the 95% CI. Locations with fewer than 30 reported COVID-19 deaths were not considered in the analysis.

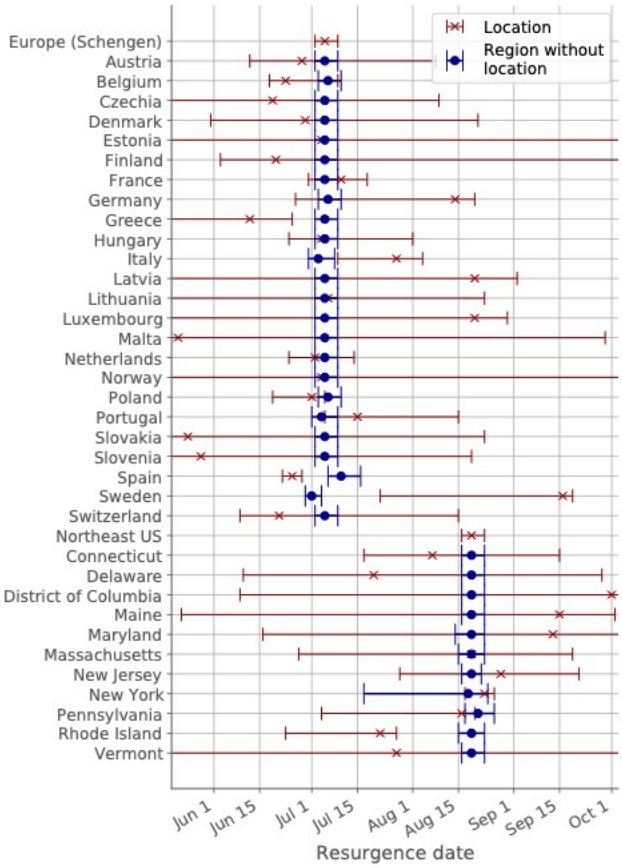

Figure 2. Timing of the COVID-19 outbreak resurgence in Europe and the US Northeast. Dates of the minimum infectious population reached are shown for each location (red crosses) and for the whole region without the population of the location indicated (blue circles). The intervals represent the 95% CI. Locations with fewer than 30 reported COVID-19 deaths were not considered in the analysis.