medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

Impact of seasonal forcing on a potential SARS-CoV-2 pandemic

Richard A. Neher,[1,2,*] Robert Dyrdak,[3,4] Valentin Druelle,[1,2] Emma B. Hodcroft,[1,2] and Jan Albert[3,4]
[1] *Biozentrum, University of Basel, Basel, Switzerland*
[2] *Swiss Institute of Bioinformatics, Basel, Switzerland*
[3] *Department of Clinical Microbiology, Karolinska University Hospital, Stockholm, Sweden*
[4] *Department of Microbiology, Tumor and Cell Biology, Karolinska Institute, Stockholm, Sweden*

(Dated: February 13, 2020)

A novel coronavirus (SARS-CoV-2) first detected in Wuhan, China, has spread rapidly since December 2019, causing over 45,000 PCR confirmed infections and more than 1,000 fatalities (as of Feb 12, 2020). Imported cases and small transmission clusters have been reported globally. Early data suggest the virus transmits readily and a pandemic cannot be ruled out. Data from routine diagnostics show a strong and consistent seasonal variation of the four endemic coronaviruses (229E, HKU1, NL63, OC43). We use these data to explore the effect of seasonal variation in transmissibility on a potential SARS-CoV-2 pandemic. A model allowing for many subpopulations of different size with variable parameters of SARS-CoV-2 spread shows how a pandemic could unfold in 2020-2022. Simulations of different scenarios show that plausible parameters result in a peak in temperate regions of the Northern Hemisphere in winter 2020/2021. A smaller range of parameters suggests a peak in the first half of 2020 or two peaks of similar magnitude. Variation in transmission and migration rates can result in substantial variation in prevalence between regions.

While the uncertainty in parameters is large, the scenarios we explore show that transient reductions in the incidence rate might be due to a combination of seasonal variation and infection control efforts but do not necessarily mean the epidemic is contained. Seasonal forcing on SARS-CoV-2 should thus be taken into account in the further monitoring of the global transmission. The likely aggregated effect of seasonal variation, infection control measures, and transmission rate variation is a prolonged pandemic wave with lower prevalence at any given time, thereby providing a window of opportunity for better preparation of health care systems.

On Jan 30, 2020, the World Health Organisation (WHO) declared the spread of a new coronavirus, SARS-CoV-2 (Gorbalenya, 2020)), as a public health emergency of international concern (WHO Emergency Committee, 2020b). The virus was first identified in patients with pneumonia in the city of Wuhan in the Hubei province, China, in December 2019 (Liangjun *et al.*, 2020). The clinical presentation of the illness caused by SARS-CoV-2, called COVID-19, appears to range from mild or asymptomatic to severe and fatal respiratory illness (WHO Emergency Committee, 2020a), but the exact spectrum of disease presentation is still unclear. The potential for global spread, i.e. a pandemic, of SARS-CoV-2 is currently not known, but the virus has spread at an alarming rate in Wuhan, the epicenter of the outbreak. Furthermore, the virus has spread to all provinces of China and small clusters of local spread have been reported from several countries, e.g. Singapore, Germany, and the UK (Rothe *et al.*, 2020; Singapore Ministry of Health, 2020; WHO Emergency Committee, 2020a).

The basic reproduction number (R_0), which describes the average number of new infections per infected SARS-CoV-2 case, has been estimated to be around $R_0 = 2 \ldots 3$ (2.2 with 90% high density interval 1.4–3.8 (Riou and Althaus, 2020) or 2.7 with a 95% CrI of 2.47–2.86 (Wu *et al.*, 2020)) Higher estimates have also been reported (Yang *et al.*, 2020). Importantly, R_0 is not a biological constant for a pathogen, but is affected by factors such as environmental conditions and the behaviour of infected individuals. One such environmental factor is climate which modulates transmissibility throughout the year. As a result, many respiratory viruses show clear seasonal variation in prevalence; the most well-known example being seasonal influenza which peaks every winter in the temperate zone of the Northern Hemisphere (Petrova and Russell, 2018). A similar pattern is seen for the four seasonal human coronaviruses: HKU1, NL63, OC43 and 229E (hereafter collectively referred to as "seasonal CoVs") (Al-Khannaq *et al.*, 2016; Friedman *et al.*, 2018; Galanti *et al.*, 2019; Góes *et al.*, 2019; Huang *et al.*, 2017; Killerby *et al.*, 2018). These viruses cause respiratory infections which usually are mild and primarily affect young children.

Previous influenza pandemics have swept the world in multiple waves often but not always coinciding with winter months in temperate climates (Amato-Gauci *et al.*, 2011; Taubenberger *et al.*, 2019; Viboud *et al.*, 2005, 2016). The 1968-1970 global influenza pandemic was sparked by a new influenza A/H3N2 virus, with a divergent hemagglutinin protein, which replaced the A/H2N2 virus circulating for 10 years previously (Viboud *et al.*,

* Correspondence to: Richard Neher, Biozentrum, Klingelbergstrasse 70, 4056, Basel, Switzerland. **richard.neher@unibas.ch**

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

2005). The virus spread rapidly, but the viral dynamics and mortality was not synchronized across countries. While many experienced two more-severe flu seasons in the winters of 1969 and 1970, the US had higher mortality in the first season, while European countries, Japan, and Australia, had higher mortality in the second, and Canada had roughly equal mortality in both (Viboud *et al.*, 2005).

The 2009 pandemic H1N1 virus (A/H1N1pdm09) originated in March 2009 in Mexico and spread around the globe within weeks. Only a few European countries saw substantial circulation of H1N1pdm09 in the spring of 2009. Instead the virus showed low prevalence over the summer and pronounced peaks in the following autumn and winter in many countries (Amato-Gauci *et al.*, 2011). A/H1N1pdm09 has subsequently transitioned into a seasonal pattern causing winter epidemics in temperate climates.

Here we use data on seasonal variation in prevalence of seasonal CoVs in Sweden and model the impact of this variation on the possible future spread of SARS-CoV-2 in the temperate zone of the Northern Hemisphere. We also explore different scenarios of SARS-CoV-2 spread in temperate and tropical regions and show how variation in epidemiological parameters affects a potential pandemic and the possibility of transitioning to an endemic state.

I. SEASONAL CORONAVIRUS PREVALENCE

Data on seasonal variation of HKU1, NL63, OC43 and 229E diagnoses in respiratory samples was obtained from the routine molecular diagnostics at the Karolinska University Hospital, Stockholm, Sweden. The laboratory provides diagnostic services to six of seven major hospitals and approximately half of outpatient care in the Stockholm county (2.2 million inhabitants). We extracted pseudonymized data on all analyses for the four viruses between Jan 1, 2010 and Dec 31, 2019. The dataset included a total of 52,158 patient samples with 190,257 diagnostic tests, of which 2,084 were positive for any of the coronaviruses (229E = 319; NL63 = 499; OC43 = 604; HKU1 = 355; OC43/HKU1 = 307). Metadata included information about date of sampling and age of patient. In the period of Jan 1, 2010 to Nov 5, 2017, the coronavirus diagnostic was done using in-house assays (Tiveljung-Lindell *et al.*, 2009). From Nov 6, 2017 to Dec 31, 2019, samples were analysed using the commercial kit Allplex Respiratory Panels (Seegene Inc., Seoul (South Korea)). This commercial kit does not distinguish between HKU1 and OC43, and for this reason positive tests for these two viruses were combined for the entire study period.

The fraction of tests that were positive for the four seasonal CoVs showed a strong and consistent seasonal variation, see Fig. 1. From December to April approximately 2% of tests were positive, while less than 0.2% of tests were positive between July to September, i.e. a 10-fold difference (Fig. 1, right). The strength of variation of the transmission rate through the year could be of high relevance to the spread of SARS-CoV-2 in 2020 and following years.

II. BASIC MODEL

We consider simple SIR models (Kermack and McKendrick, 1991) of the form

$$\begin{aligned} \frac{d}{dt}S &= b(1-S) - \beta(t)SI \\ \frac{d}{dt}I &= -(\nu + b + e)I + \beta(t)SI + i \\ R &= 1 - S - I \end{aligned} \tag{1}$$

where $\beta(t)$ is the rate at which an infected individual infects a susceptible one, ν is the recovery rate, b is the population turn-over rate, e is the emigration rate, and i is the influx of infected individuals. The population turnover rate is immaterial for a pandemic scenario, but important for our analysis of seasonal CoV, and should be interpreted as the sum of the birth rate and the rate at which previously immune individuals become susceptible due to immune waning and escape. We review general properties of such model in the Supplementary Materials. Following previous work, we parameterize transmissibility as

$$\beta(t) = \beta_0 \left(1 + \epsilon \cos(2\pi(t-\theta))\right) \tag{2}$$

where β_0 is the average annual infection rate, ϵ is the amplitude of seasonal forcing which modulates transmissibility through the year, and θ is the time of peak transmissibility (Chen and Epureanu, 2017; Dushoff *et al.*, 2004). For simulations of the pandemic, we will add an additional term to $\beta(t)$ that accounts for infection control measures in heavily affected areas, see Supplementary Materials.

III. MODEL PARAMETERIZATION USING SEASONAL COV OBSERVATIONS

Seasonal CoVs are endemic throughout the world and we therefore expect that viruses are imported throughout the year. We model this import as a constant influx i. Furthermore, we assume that humans suffer from a seasonal CoV infection on average every 10 years ($b = 0.1/y$), an $R_0 = 2.5$, and a recovery rate (time between infections) of 10 days.

With these assumptions, we can solve the model and compare the resulting trajectories to the seasonal variation in prevalence of seasonal CoVs, see Fig. 2. We

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

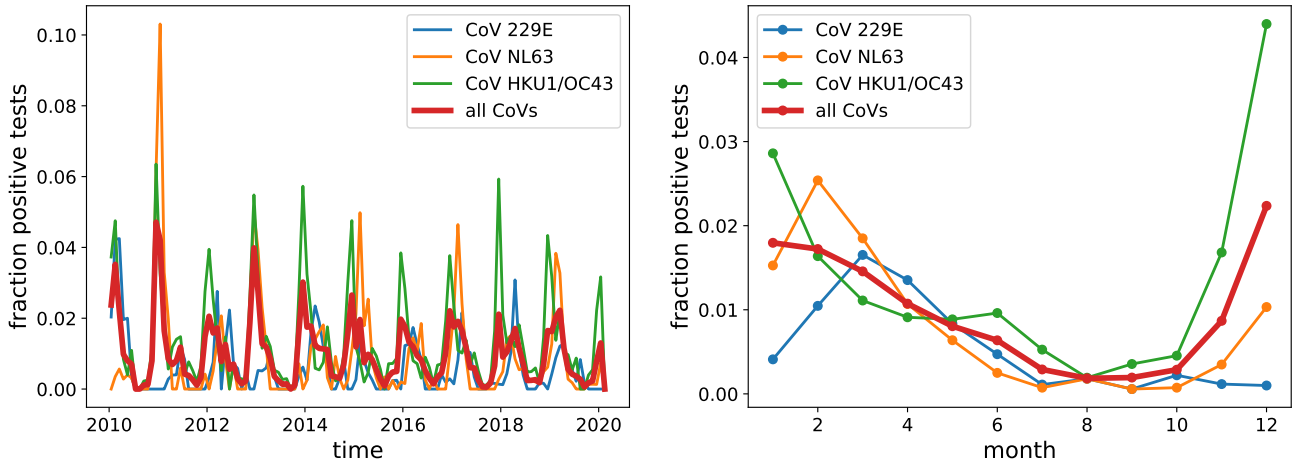

FIG. 1 **Seasonal variation in the fraction of positive CoV tests in Stockholm, Sweden.** The left panel shows test results between 2010 and 2019. The right panel shows aggregated data for all years. All CoVs show a marked decline in summer and autumn, with HKU1/OC43 peaking January–December, and NL63 and 229E peaking in February–March.

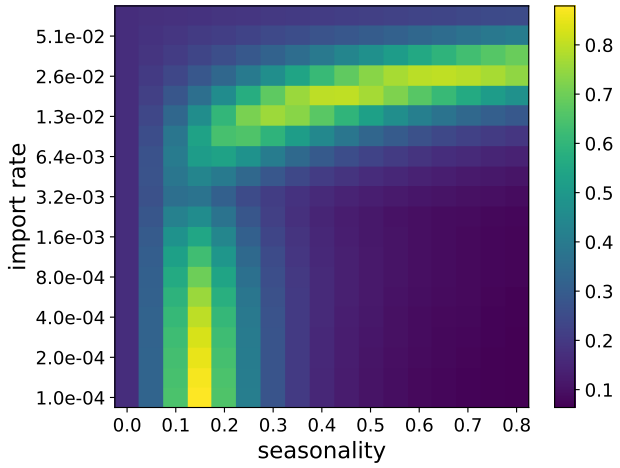

FIG. 2 **Compatibility of SIR model trajectory with observations.** The heatmap shows the inverse mean squared deviation between the model trajectories and the observed seasonal forcing in seasonal CoV prevalence. Model and observations are compatible (yellow shading) in a region of parameter values corresponding to low migration/weak seasonality and second region at high migration/strong seasonality.

assess fit quality by calculating the squared deviation of observed and predicted prevalence relative to their respective mean values. Simulations of the model are compatible with observations in two separate regions of parameter space: If Northern Europe was very isolated with less than 1 in 1,000 individuals returning with a seasonal CoV infection from abroad, weak seasonality of around $\epsilon = 0.15$ would be sufficient to generate strong variation through the year compatible with observations (Fig. 2, bottom-left ridge). In this regime, prevalence is oscillating intrinsically with a period that is commensurate with annual seasonal oscillations giving rise to a resonance phenomenon with annual or biennial patterns even for weak seasonal forcing (Chen and Epureanu, 2017; Dushoff *et al.*, 2004).

If the rate of import of seasonal CoV infections is higher, imports dampen the resonance and much stronger seasonality, with values between $\epsilon = 0.3$ and 0.7, is required to fit the observations (Fig. 2, top-right-and-center ridge). In this regime, seasonal variation in transmissibility modulates the size of micro-outbreaks triggered by imported cases in a mostly immune population.

These two scenarios differ slightly in the time of year at which peak transmissibility θ occurs: When transmission is mostly local and seasonality is amplified by resonance, θ needs to be around October–November to fit the data with most cases in December–January. In the second scenario with high connectivity, θ needs to be in December–January coinciding with the peak in prevalence. Given that most countries are highly connected, we focus here on exploring the high-import and strong seasonal forcing scenario. This scenario, with maximal β in mid-winter, is also more compatible with climate variation around the year.

IV. SCENARIOS FOR SARS-COV-2 PANDEMICS IN 2020 AND 2021

The analysis of seasonal CoV prevalence patterns allowed us to constrain parameter ranges and explore different scenarios of SARS-CoV-2 spread around the globe, in particular in temperate climates like Northern Europe. Here we explore scenarios where temperate regions have a seasonal forcing of between $\epsilon = 0.3$ and 0.7 and migration rates of 0.01/year. Early estimates suggest an incubation time of about 5 days (Backer *et al.*, 2020) and an

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

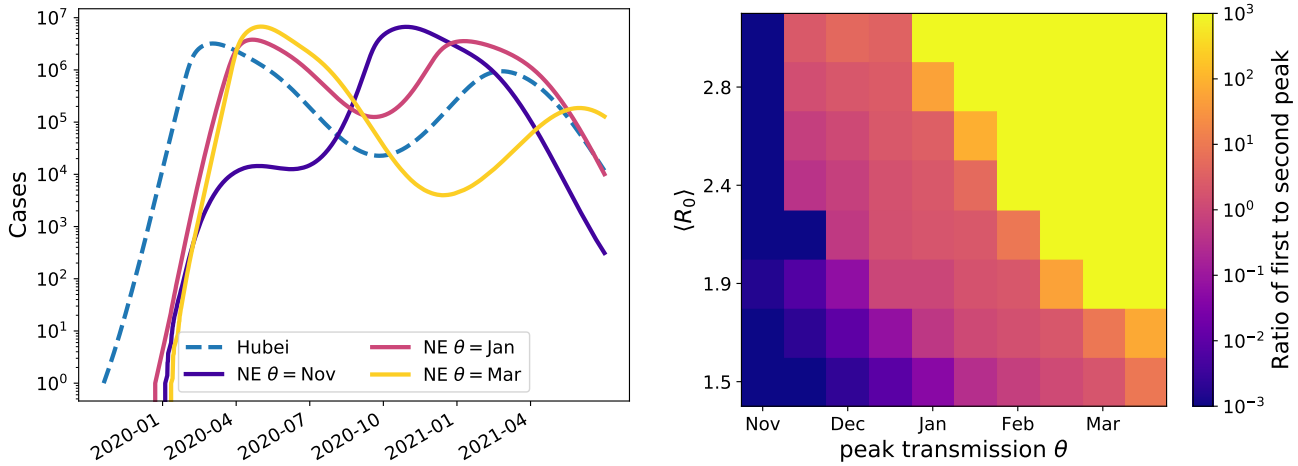

FIG. 3 **Model predictions for SARS-CoV-2 case numbers in temperate zones for a pandemic scenario**. The left panel shows example trajectories assuming SARS-CoV-2 transmissibility peaks in November, January, or March. These outbreaks in Northern Europe ('NE') are assumed to be seeded by the outbreak in Hubei (model trajectory shown as a dashed line). Within the model, these cases are exported at rate of 0.01/year to temperate Northern Europe with an average $\langle R_0 \rangle = 1.8$ and seasonal forcing of $\epsilon = 0.5$. Corresponding graphs for different values of $\langle R_0 \rangle$ and the migration rate are shown in Supplemental Fig. 3. The right panel shows the ratio of the first and second peak for a range of different combinations of R_0 and θ. The yellow area corresponds to parameter combinations with essentially only an early peak similar to the yellow line on the left. The blue/purple area shows parameter combinations for which a peak in late 2020 dominates, as with the purple line on the left, while the central pink/orange band shows the combinations giving rise to two comparable peaks. These simulations are for $\epsilon = 0.5$. Similar results were obtained for $\epsilon = 0.3$ and 0.7, see Supplementary Fig. 4.

average serial interval of 7-8 days (Wu *et al.*, 2020), suggesting rapid transmission once symptoms set in. Our model does not differentiate between symptomatic and asymptomatic cases. Furthermore, the model does not account for the incubation time and therefore generates too many early transmissions. We therefore use a slightly lower recovery rate of $\nu = 0.1/day$. To match the R_0 estimates for the early outbreak with our parameterization of transmissibility in Eq. 2 we need to account for the fact that December/January are winter months in Hubei and peak transmissibility in Hubei likely corresponds to $\theta \approx 0$ (0 being the beginning of the year, so a θ in December/January). An $R_0 \approx 2.5$ in winter in Hubei and a seasonal forcing of $\epsilon = 0.4$ implies an annual average $\langle R_0 \rangle = \beta_0/\nu = 1.8$. This reasoning leads to our parameter choice of $\beta_0 = 65$/year, $\nu = 36$/year, $\theta = 0$. We assume the outbreak started at $t = 2019.8$ in Hubei with one infected individual and use $N = 6 \times 10^7$ as population size. To incorporate infection control measures, transmissibility is reduced by 50% once prevalence reached 3% (third order Hill-function, see Supplemental Material). Introductions to a location like Northern Europe with $\epsilon = 0.5$ (*i.e.* slightly stronger seasonal forcing then Hubei) are assumed to happen at a rate of 0.01 per year for each infected individual elsewhere. The simulation of the SIR model in different regions is deterministic, but migration is implemented stochastically by Poisson resampling of the average number of migrating individuals. Fig. 3 shows simulated trajectories of SARS-CoV-2 prevalence in the temperate Northern Hemisphere assuming the outbreak started in Hubei early December 2019. Depending whether the peak transmissibility of SARS-CoV-2 in the northern temperate zone is in November, January, or March, the simulation predicts a main peak in the first half of 2020, a main peak in winter 2020/2021, or two similarly sized peaks.

To explore possible scenarios more systematically, we ran such simulations for a range of values for R_0 and peak transmissibility θ and recorded whether we observe and early peak, a late peak, or a two peaks. The right panel of Fig. 3 shows the ratio of the height of these peaks for different values. Rapid growth (high R_0) and late transmission peaks result in a large peak in the first half 2020, while lower R_0 and transmission peaks in early winter favor a large secondary peak. These two scenarios are separated by a band of parameter values that give rise to two pandemic waves in the winters of 2020 and 2021 in the Northern Hemisphere.

The uncertainty in parameter values and the potential impact of infection control measures imply that all scenarios are plausible and should be considered when developing pandemic prevention and containment strategies.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

V. GLOBAL PROJECTIONS

Outbreaks grow exponentially within well-mixed communities, and at a certain rate the virus will be carried to other regions and potentially seed new outbreaks. Such export to new locations is initially unlikely, but becomes next to certain once the outbreak size exceeds the inverse probability that a given individual migrates while infected.

Every location has a different socio-economic profile such that the growth rate of the epidemic (and hence R_0) might differ. The superposition of many such subpopulations with a range of $\langle R_0 \rangle$ values and seasonal variation in transmission will result in dynamics that are qualitatively different from a single population SIR model. In particular, such variation result in a pandemic spread out over 2 years before the virus possibly becomes endemic.

Fig. 4 shows the result of such a simulation of 1,000 populations. Populations were divided between northern temperate (50%), southern temperate (10%), and tropical (40%) and assigned parameters as follows:

- $\langle R_0 \rangle$ was drawn from a normal distribution with mean 1.8 (see Supplementary Fig. 5 for $\langle R_0 \rangle = 1.4$ and 2.7) and standard deviation 0.5.
- Seasonal forcing ϵ was drawn from a uniform distribution between 0.25 and 0.75 for temperate regions, between 0 and 0.2 for tropical regions.
- Peak transmissibility θ of temperate regions was drawn from normal distributions with standard deviation 0.1 and peak at 0 for northern regions and 0.5 for southern regions. θ for tropical regions was chosen uniformly from between 0 and 1.
- Population sizes were drawn from a log-normal distribution with $\sigma = 1$ and a mean such that all populations sum to 7.6 billion.
- Migration rates were sampled from a log-normal distribution with $\sigma = 1$ and a mean of 0.01.

For Hubei, we use the same parameters as described in section IV.

The variation in R_0 and migration rate result in a super-position of fast and slow epidemics seeded at different times. The initial phase is dominated by fast epidemics driving rapid dispersal, in particular in the tropics, while slow epidemics dominate later in 2020 and 2021. With the parameter setting used in Fig. 4, the Northern temperate regions see most circulation in winter 2020/2021. In accordance with Fig. 3, this peak shifts more towards early 2020 for higher R_0, see Supplementary Fig. 5.

After several years, SARS-CoV-2 could become a seasonal CoV with characteristic winter outbreaks as shown in Fig. 1. Such a scenario is demonstrated in Fig. 5 where a simulation similar to the one shown in Fig. 4 is run for 12 years, with the added assumption that after infection an individual become susceptible to SARS-CoV-2 again at a rate of 0.1 per year as we assumed for seasonal CoV above. After a pronounced low in 2020-2024, prevalence recovers and settles into a seasonal pattern, similar to that of the four existing seasonal CoVs.

VI. DISCUSSION

We report on the possible influence of seasonal variation on the spread of SARS-CoV-2 in the Northern Hemisphere, should the current outbreak progress to a pandemic. We find that seasonal variation in transmissibility has the potential to modulate the spread of SARS-CoV-2 with a wide range of possible outcomes that need to be taken into account when interpreting case counts and projecting the outbreak dynamics. The onset of spring and summer could, for example, give the impression that SARS-CoV-2 has been successfully contained, only for infections to increase again in 2020-2021 winter season. Even in Hubei virus circulation might decrease due to arrival of spring and summer but might increase again towards the end of the year. Whether a pandemic in the temperate regions of the Northern Hemisphere would peak early in 2020, late in 2020, or show multiple waves as H1N1pdm did in 2009, depends on the timing of peak transmissibility and the rate of spread (R_0 and serial interval).

This study is meant as an exploration of how such a pandemic could unfold, not as a prediction of any particular scenario. The results we present are critically dependent on the assumptions i) that the outbreak will develop into a pandemic, ii) that the transmissibility of SARS-CoV-2 shows seasonal variability of sufficient strength (range $\epsilon = 0.3$ to 0.7), and iii) that parameters like R_0 estimated from the early phase of the outbreak are comparable in other populations.

These assumptions are not implausible but not certain: cases of SARS-CoV-2 has been detected in 24 countries apart from China (WHO Emergency Committee, 2020a), and mild or asymptomatic cases make detection and thus prevention of spread by isolation challenging, *e.g.* airport screening as a preventive measure is unlikely to prevent spread and local seeding (Quilty *et al.*, 2020). Person-to-person transmission of the virus has been documented in several countries outside of China. It is likely that not all exported cases have been detected, and some may have seeded outbreaks outside of China that have yet to be detected. It is estimated that if SARS-CoV-2 continues to spread in China, and particularly if the areas of Beijing and Shanghai are affected, the risk of importation of cases to Europe increases (Pullano *et al.*, 2020).

The seasonal CoVs show a strong and consistent sea-

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

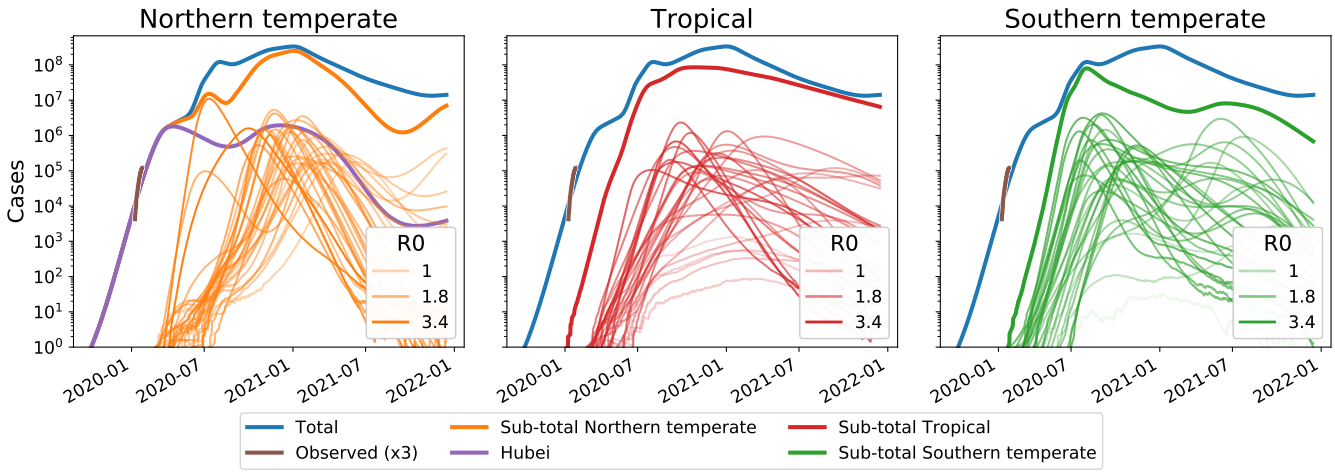

FIG. 4 **Extended circulation through overlapping epidemics in variable subpopulations.** These simulations of a pandemic scenario assume 1,000 sub-populations with an average $\langle R_0 \rangle$ of 1.8 and standard deviation 0.5, 40% of which have weak seasonal forcing $\epsilon \in [0, 0.2]$ (tropical) and the remainder have strong variation with $\epsilon \in [0.25, 0.75]$. The super-position of many variable epidemics can result in a global prevalence that decays only slowly through 2020 and 2021. Lighter lines have lower R_0, darker lines have higher R_0. The actual observed case counts reported for Hubei are added (brown line) and multiplied by three to account for possible under-reporting of mild cases. A subset of 30 randomly chosen simulations are plotted for each region. Analogous figures for different R_0 parameter values are shown in Supplementary Fig. 5.

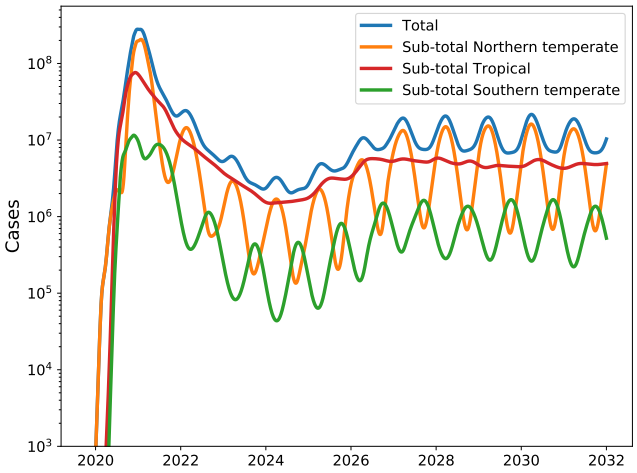

FIG. 5 **Transition to an endemic seasonal virus.** If previously infected individuals can be reinfected after some time, as for example by seasonal influenza virus, SARS-CoV-2 could develop into a seasonal CoV that returns every winter. This would typically happen at much lower prevalence than peak pandemic levels. This graph assumes reinfection on average every 10 years.

sonal variation, and modeling suggests that this requires strong variation in transmissibility throughout the year. It should be noted, however, that SARS-CoV-2 does seem to transmit in tropical climates like Singapore, and so winter is not a necessary condition of SARS-CoV-2 spread. Precise values for the underlying model parameters and the effect of infection control measures are currently unavailable. For this reason we explored a range of parameter values to assess the robustness of the results to model assumptions.

The simulations presented here are scenarios that emerge from simplified abstract models, but they nevertheless demonstrate that a wide variety of outcomes are compatible within the limits of the current knowledge about the outbreak. The implications of our work are that: 1) reductions in prevalence need not be attributable to successful interventions, but could be due to seasonal variation in transmissibility, 2) sub-population dynamics can differ greatly, meaning that case count trajectories in one country should be used cautiously to inform projections in a second country, even in the same climate zone, 3) seasonal variation might slow down a pandemic and thereby provide a window of opportunity for better preparation of health care systems world-wide by scaling up capacity for care and diagnostics, and potentially through rapid development of antivirals and vaccine, and 4) after several years SARS-CoV-2 could develop into an endemic seasonal CoV similar to the transition of the 2009 A/H1N1 pandemic influenza virus into a seasonal influenza virus.

The overall impact of a potential SARS-CoV-2 pandemic depends critically on the case fatality ratio (CFR), which we have not modelled here. At present, uncertainty in the CFR is high due to likely over-representation of severe cases in the statistics and a delay between diagnosis and recovery/death (Battegay *et al.*, 2020). Even with this unknown, seasonal variation in transmissibility of SARS-CoV-2 and underlying differences in migration, introduction times, and attack rate should thus be taken into account when monitoring and projecting

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

global transmission, planning further surveillance of the epidemic, and developing pandemic prevention and containment strategies.

VII. CODE AND DATA AVAILABILITY

All relevant data and script that generate the graphs are available in a dedicated github repository at `github.com/neherlab/CoV_seasonality`.

ACKNOWLEDGEMENTS

We are grateful to all physicians, scientists, and public health authorities globally for sharing case numbers and early research findings without delay. We would like to thank Christian Althaus for critically reading the manuscript and providing timely and valuable feed back. This research was supported in part by the National Science Foundation under Grant No. NSF PHY-1748958 (RAN during his visit to KITP).

AUTHORS' CONTRIBUTIONS

JA conceived the idea of the study. RAN developed the model and simulations. JA and RD provided the data on seasonal CoVs. RAN, EBH and VD performed the modeling and simulation, and generated the graphs. All authors interpreted the data, drafted and commented on the manuscript, and contributed to the final version.

REFERENCES

Al-Khannaq, M. N., K. T. Ng, X. Y. Oong, Y. K. Pang, Y. Takebe, J. B. Chook, N. S. Hanafi, A. Kamarulzaman, and K. K. Tee (2016), The American journal of tropical medicine and hygiene **94** (5), 1058.

Amato-Gauci, A., P. Zucs, R. Snacken, B. Ciancio, V. Lopez, E. Broberg, P. Penttinen, A. Nicoll, and c. o. b. o. t. E. I. S. Network (EISN) (2011), Eurosurveillance **16** (26), 19903.

Backer, J. A., D. Klinkenberg, and J. Wallinga (2020), medRxiv , 2020.01.27.20018986.

Battegay, M., R. Kuehl, S. Tschudin-Sutter, H. H. Hirsch, A. F. Widmer, and R. A. Neher (2020), Swiss Medical Weekly **150** (0506), 10.4414/smw.2020.20203.

Chen, S., and B. Epureanu (2017), Journal of theoretical biology **415**, 137.

Dushoff, J., J. B. Plotkin, S. A. Levin, and D. J. D. Earn (2004), Proceedings of the National Academy of Sciences **101** (48), 16915.

Friedman, N., H. Alter, M. Hindiyeh, E. Mendelson, Y. Shemer Avni, and M. Mandelboim (2018), Viruses **10** (10), 515.

Galanti, M., R. Birger, M. Ud-Dean, I. Filip, H. Morita, D. Comito, S. Anthony, G. A. Freyer, S. Ibrahim, B. Lane, *et al.* (2019), Influenza and Other Respiratory Viruses **13** (3), 226.

Góes, L. G. B., R. M. Zerbinati, A. F. Tateno, A. V. de Souza, F. Ebach, V. M. Corman, C. A. Moreira-Filho, E. L. Durigon, L. V. R. F. da Silva Filho, and J. F. Drexler (2019), Journal of medical virology.

Gorbalenya, A. E. (2020), bioRxiv , 2020.02.07.937862.

Huang, S.-H., M.-C. Su, N. Tien, C.-J. Huang, Y.-C. Lan, C.-S. Lin, C.-H. Chen, and C.-W. Lin (2017), Journal of Microbiology, Immunology and Infection **50** (6), 763.

Kermack, W. O., and A. G. McKendrick (1991), Bulletin of Mathematical Biology **53** (1), 33.

Killerby, M. E., H. M. Biggs, A. Haynes, R. M. Dahl, D. Mustaquim, S. I. Gerber, and J. T. Watson (2018), Journal of Clinical Virology **101**, 52.

Liangjun, C., L. Weiyong, Z. Qi, X. Ke, Y. Guangming, W. Weichen, S. Ziyong, L. Fang, W. Kailang, Z. Bo, M. Yi, Z. Wenxia, C. Yu, L. Yirong, S. Mang, L. Ke, and L. Yingle (2020), Emerging Microbes & Infections **9** (1), 313.

Petrova, V. N., and C. A. Russell (2018), Nature Reviews Microbiology **16** (1), 47.

Pullano, G., F. Pinotti, E. Valdano, P. Y. Boelle, C. Poletto, and V. Colizza (2020), Euro Surveill **25** (4), 10.2807/1560-7917.Es.2020.25.4.2000057.

Quilty, B. J., S. Clifford, S. Flasche, R. M. Eggo, C. nCoV working group, *et al.* (2020), Eurosurveillance **25** (5), 2000080.

Riou, J., and C. L. Althaus (2020), Euro Surveill **25** (4), 10.2807/1560-7917.Es.2020.25.4.2000058.

Rothe, C., M. Schunk, P. Sothmann, G. Bretzel, G. Froeschl, C. Wallrauch, T. Zimmer, V. Thiel, C. Janke, W. Guggemos, *et al.* (2020), New England Journal of Medicine.

Singapore Ministry of Health, (2020), "Updates on 2019 novel coronavirus (2019-ncov) local situation," .

Taubenberger, J. K., J. C. Kash, and D. M. Morens (2019), Science Translational Medicine **11** (502), eaau5485.

Tiveljung-Lindell, A., M. Rotzén-Östlund, S. Gupta, R. Ullstrand, L. Grillner, B. Zweygberg-Wirgart, and T. Allander (2009), Journal of medical virology **81** (1), 167.

Viboud, C., R. F. Grais, B. A. P. Lafont, M. A. Miller, and L. Simonsen (2005), The Journal of Infectious Diseases **192** (2), 233.

Viboud, C., L. Simonsen, R. Fuentes, J. Flores, M. A. Miller, and G. Chowell (2016), The Journal of Infectious Diseases **213** (5), 738.

WHO Emergency Committee, (2020a), "Novel Coronavirus (2019-nCoV) Situation Report – 23," .

WHO Emergency Committee, (2020b), "Statement on the second meeting of the international health regulations (2005) emergency committee regarding the outbreak of novel coronavirus (2019-ncov)," .

Wu, J. T., K. Leung, and G. M. Leung (2020), The Lancet 10.1016/S0140-6736(20)30260-9.

Yang, Y., Q. Lu, M. Liu, Y. Wang, A. Zhang, N. Jalali, N. Dean, I. Longini, M. E. Halloran, B. Xu, X. Zhang, L. Wang, W. Liu, and L. Fang (2020), medRxiv , 2020.02.10.20021675.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

SUPPLEMENTARY INFORMATION

The SIR model introduced in equations 1 describes the fraction of susceptible $S(t)$ and infected individuals $I(t)$ in a population. The fraction of recovered individuals R is $R = 1 - S - I$. A simplified scheme of the model is presented in figure 1 with constant transmissibility ($\epsilon = 0$ in equation 2) and without emigration/influx ($e = 0$ and $i = 0$). The simplified system has a steady state

$$S_0 = \frac{\nu + b}{\beta} = \frac{1}{R_0}$$
$$I_0 = \frac{b(R_0 - 1)}{\beta} \tag{3}$$

This steady state exists when the parameter $R_0 > 1$. For cases where $R_0 < 1$ the epidemic dies out as each infected person generate less than one subsequent infection.

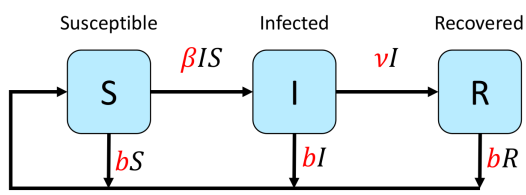

Figure S 1 Illustration of the SIR model used with its parameters.

For our analysis of seasonal CoVs, it is instructive to study the stability of this steady state and show that the quantities $[S, I]$ tend to show damped oscillations around the fixed points in Eq. 3 with a period

$$T = \frac{4\pi}{\sqrt{4b(\nu(R_0 - 1) - b) - R_0^2 b^2}} \tag{4}$$

This behavior is intrinsic to the system and holds whenever transmissibility is constant.

SIR models can show resonance phenomena when the period is about one or two years (Chen and Epureanu, 2017; Dushoff *et al.*, 2004). Resonance is less likely when the period is much longer than the the period of annual forcing. Fig. 2 shows the intrinsic period for the parameter range considered. In the center of the graph the intrinsic period is around 3 years, suggesting that resonance is not strong.

In a fully susceptible population, the number of infected individuals initially grows as $I(t) \sim e^{(\beta-\nu-b)t} = e^{(R_0-1)t/(\nu+b)}$. The doubling time is $\tau_2 = (\nu + b)\log(2)/(R_0 - 1)$. Since $b \ll \nu$, we can safely neglect b in the early phase of the pandemic, but it is important when interpreting endemic seasonal CoV data.

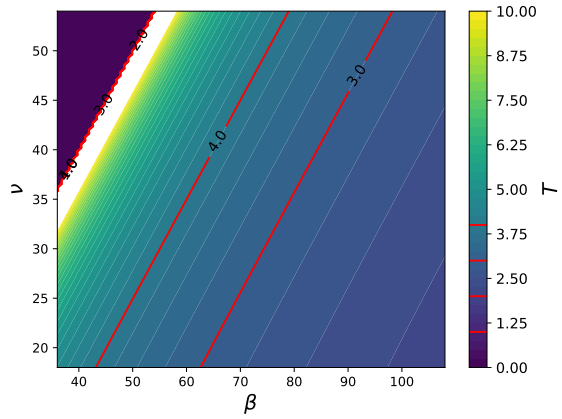

Figure S 2 Length of the intrinsic period of the SIR model for β and ν values in the range explored in this paper. The duration of immunity is assumed to be 10 years. The dark blue region corresponds to $R_0 < 1$.

Migration between subpopulations

The term i and e in equation 1 represent the influx and emigration in a population which we model as an exchange of infected individuals while ignoring the migration of susceptible or recovered individuals. Such migration is operating in Figures 3, 4 and 5. A proportion e of the infected people are removed due to emigration while an influx i is added. The influx i of each population is computed as the product of the total number of infected people in all population, the migration rate of the population in question, and its size. In that sense, the infected persons form a global "reservoir" from which each population takes its influx.

Effect of infection control measures

The SIR model as formulated in Eq. 1 accounts for seasonality but otherwise assumes the transmission dynamics is constant. In reality, individuals change their behavior when they are aware of a contagious disease. This in particularly true for the current SARS-CoV-2 outbreak where tens of millions of people are under quarantine measures and travel restrictions. In order to account for such behavioral changes, we add a factor H to equation 1 that modulates transmission:

$$\frac{d}{dt}I = -(\nu + b + e)I + (1 - H)\beta SI + i \tag{5}$$

where $H = c \cdot \frac{I^3}{(K^3+I^3)}$ with c the containment value. The term H is a Hill function of order 3 and inflection point K. We use this to model a decrease in the rate of new infection smoothly decrease by a factor c. We used $c = 0.5$ and $K = 0.03$, implying that once a disease is at a prevalence of 3%, containment measures reduce transmission

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

by 50%. The effect of prevention and control is difficult to assess at present and will certainly vary in time and space.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

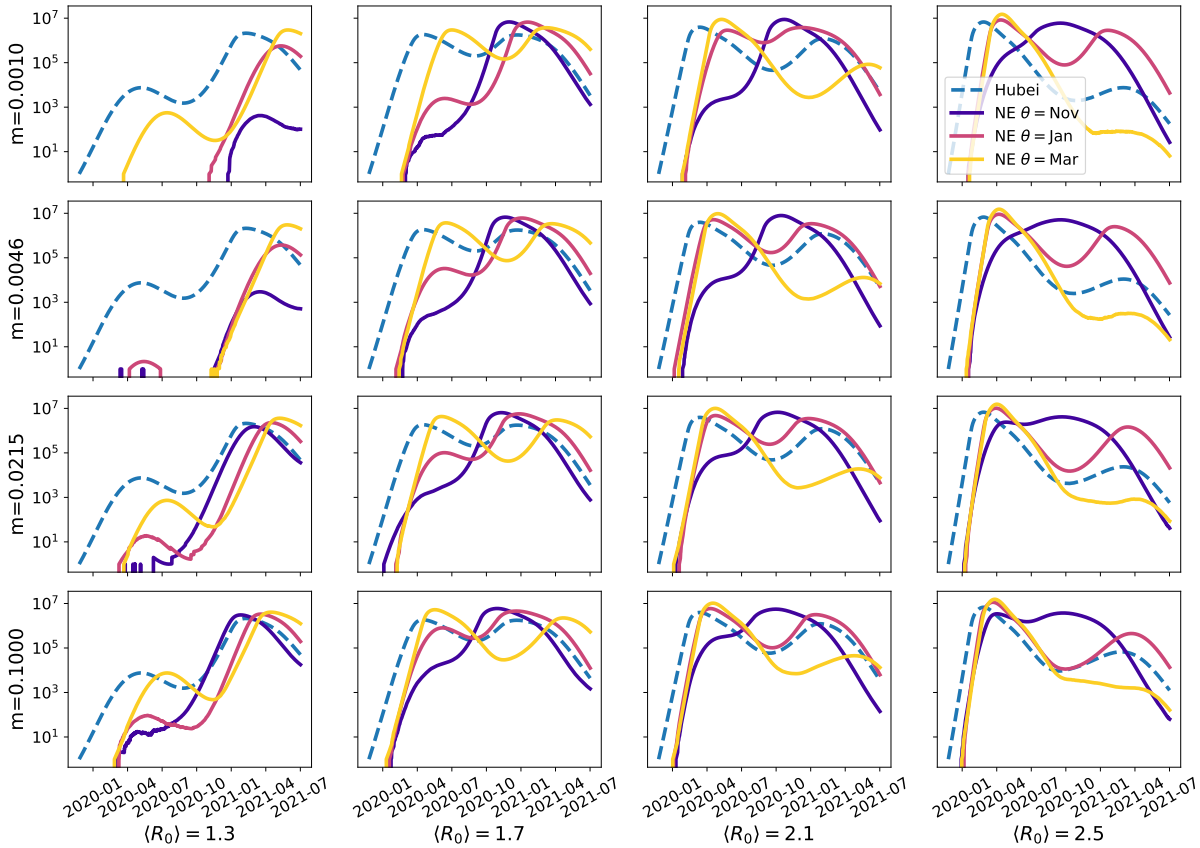

Figure S 3 **Model predictions for SARS-CoV-2 case numbers in temperate zones for a pandemic scenario, with varying R_0 and Migration.** This is the same plot as shown on the left of Figure 3, however here $\langle R_0 \rangle$ varies from 1.3 - 2.5 and Migration varies from 0.1% to 10% per year. Higher migration rates (towards the bottom of the figure) result in earlier introductions and higher likelihood of a peak in early 2020. Similarly, higher $\langle R_0 \rangle$ (towards the right of the figure) results in more rapid growth and a higher likelihood of an early peak.

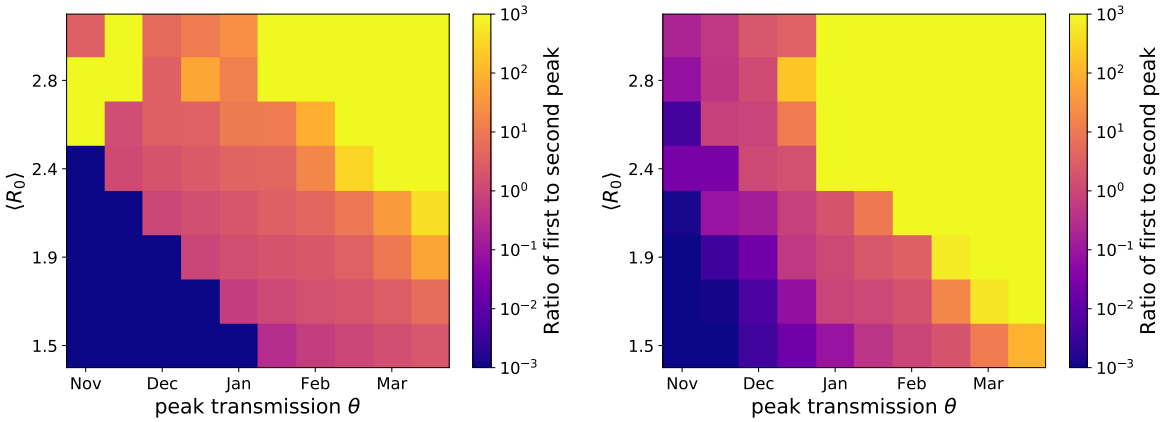

Figure S 4 **Ratio of first and second peak in temperate zone pandemic scenarios, with varying seasonal forcing.** This figure is the same as plotted on the right of Figure 3, but here seasonal forcing (ϵ) is shown at $\epsilon = 0.3$ (left) and $\epsilon = 0.7$ (right). Figure 3 uses $\epsilon = 0.5$.

medRxiv preprint doi: https://doi.org/10.1101/2020.02.13.20022806; this version posted February 17, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC 4.0 International license .

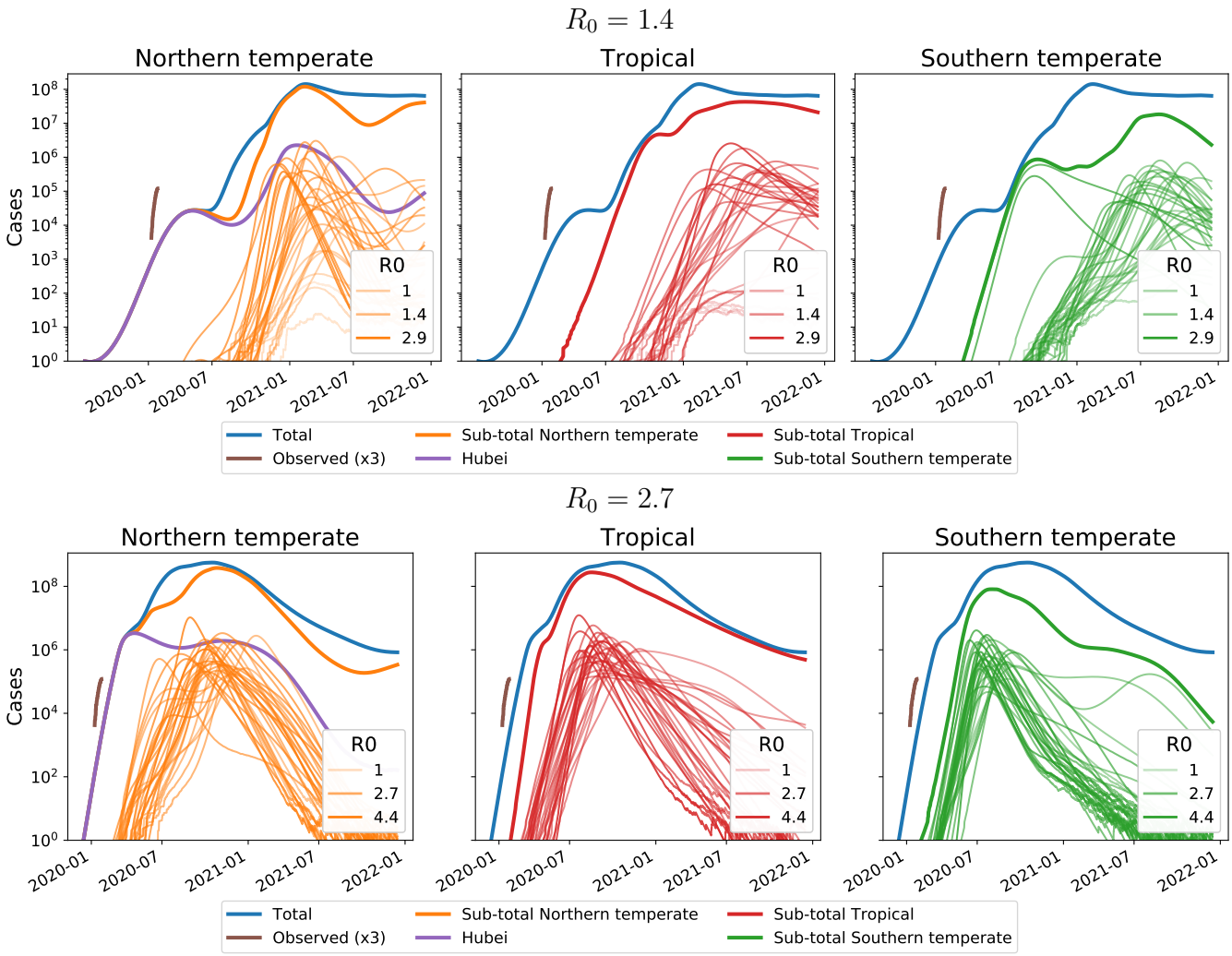

Figure S 5 **Extended circulation through overlapping epidemics in variable subpopulations, with varying R_0 values.** The same parameter values as in Figure 4 are used here, except the average $\langle R_0 \rangle$ is 1.4 (top) and 2.7 (bottom). The lower R_0 shifts the peak in case numbers to late 2020 and early 2021, while a higher R_0 leads to a peak in early-mid 2020.

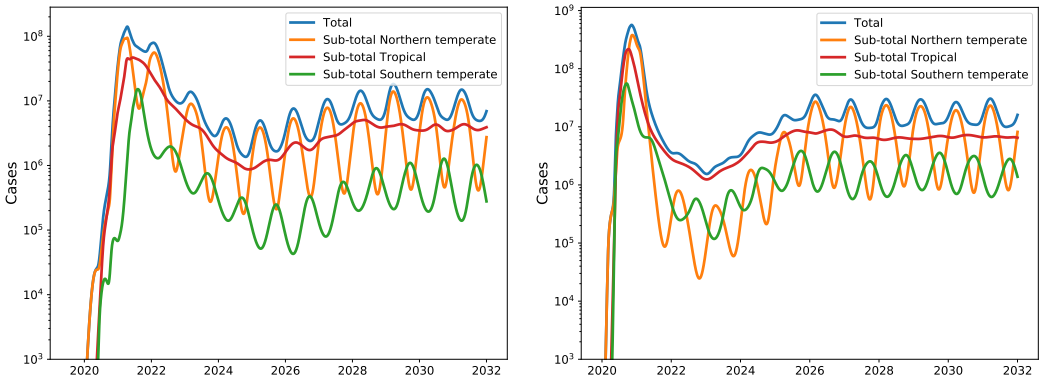

Figure S 6 **Transition to an endemic seasonal virus, with varying R_0 values.** These figures are plotted in the same manner as Figure 5, but with R_0 values of 1.5 (left) and 2.7 (right). A lower R_0 allows for a relatively smoother transition to seasonality, particularly for Northern temperate regions, while a higher R_0 leads to a longer period of comparatively depressed circulation before the establishment of stable seasonality.