RESEARCH ARTICLE

Seasonal variation in SARS-CoV-2 transmission in temperate climates: A Bayesian modelling study in 143 European regions

Tomáš Gavenčiak[1], **Joshua Teperowski Monrad**[2,3,4]*, **Gavin Leech**[5], **Mrinank Sharma**[2,6,7], **Sören Mindermann**[8], **Samir Bhatt**[9,10], **Jan Brauner**[2,8], **Jan Kulveit**[1,2]

1 Centre for Theoretical Studies, Charles University, Prague, Czech Republic, **2** Future of Humanity Institute, University of Oxford, Oxford, United Kingdom, **3** Faculty of Public Health and Policy, London School of Hygiene and Tropical Medicine, London, United Kingdom, **4** Department of Health Policy, London School of Economics and Political Science, London, United Kingdom, **5** Department of Computer Science, University of Bristol, Bristol, United Kingdom, **6** Department of Statistics, University of Oxford, Oxford, United Kingdom, **7** Department of Engineering Science, University of Oxford, Oxford, United Kingdom, **8** Oxford Applied and Theoretical Machine Learning (OATML) Group, Department of Computer Science, University of Oxford, Oxford, United Kingdom, **9** Faculty of Medicine, School of Public Health, Imperial College London, London, United Kingdom, **10** Department of Public Health, University of Copenhagen, Copenhagen, Denmark

☯ These authors contributed equally to this work.
* joshua.monrad@gmail.com

Abstract

Although seasonal variation has a known influence on the transmission of several respiratory viral infections, its role in SARS-CoV-2 transmission remains unclear. While there is a sizable and growing literature on environmental drivers of COVID-19 transmission, recent reviews have highlighted conflicting and inconclusive findings. This indeterminacy partly owes to the fact that seasonal variation relates to viral transmission by a complicated web of causal pathways, including many interacting biological and behavioural factors. Since analyses of specific factors cannot determine the aggregate strength of seasonal forcing, we sidestep the challenge of disentangling various possible causal paths in favor of a holistic approach. We model seasonality as a sinusoidal variation in transmission and infer a single Bayesian estimate of the overall seasonal effect. By extending two state-of-the-art models of non-pharmaceutical intervention (NPI) effects and their datasets covering 143 regions in temperate Europe, we are able to adjust our estimates for the role of both NPIs and mobility patterns in reducing transmission. We find strong seasonal patterns, consistent with a reduction in the time-varying reproduction number $R(t)$ (the expected number of new infections generated by an infectious individual at time t) of 42.1% (95% CI: 24.7%—53.4%) from the peak of winter to the peak of summer. These results imply that the seasonality of SARS-CoV-2 transmission is comparable in magnitude to the most effective individual NPIs but less than the combined effect of multiple interventions.

Author Summary

Building on two state-of-the-art observational models and datasets, we adapt a fully Bayesian method for estimating the association between seasonality and SARS-CoV-2 transmission in 143 temperate European regions. This approach overcomes limitations of previous analyses that do not account for the implementation of non-pharmaceutical interventions (NPIs) or mobility during the first year of the pandemic and hence may yield biased estimates of seasonal effects. We find that the seasonality of SARS-CoV-2 transmission is comparable in magnitude to the most effective individual NPIs but less than the combined effect of multiple interventions. Our findings provide valuable insights for long-term modelling and policy planning. As seasons change, it is vital that policymakers employ accurate estimates of seasonal effects. In the summer, reductions in transmission that owe to seasonality should not be misattributed to population immunity. In the winter, policymakers must avoid anticipating a greater reduction due to seasonality than will actually occur.

1 Introduction

Since the onset of the COVID-19 pandemic, the role of seasonal variation in SARS-CoV-2 transmission has received significant scientific and political attention [1]. Understanding seasonal patterns is vital, as it enables more accurate inferences about current trends in transmission and how they may change over the longer term. For example, a proper understanding of seasonality can help policymakers avoid attributing declining incidence over the summer to population immunity alone, when in fact seasonality may be playing a meaningful role.

While seasonal variation is well-established for many respiratory viral infections [2], and some studies have suggested associations between temperature, humidity, and COVID-19 incidence [3–7], other analyses have failed to show a robust role of climate and weather [8, 9], particularly when population immunity is low [10]. A recent review found that the evidence remains inconclusive [11].

A further complication is that temperature, humidity, and UV radiation plausibly affect transmission and incidence through a range of biological and epidemiological mechanisms [2, 12]. These include virus stability and viability [13, 14], host susceptibility and immune response [15, 16], human behaviour [17, 18], and social factors such as holidays and school calendars [19, 20]. This multitude of plausible causal pathways makes it exceedingly difficult to disentangle the influence of various seasonal factors, particularly given the extensive multi-collinearities and interactions between environmental, biological, and behavioural elements [21, 22]. S1 Fig shows an overview of the various causal pathways, including existing literature and evidence on the collinearities between various factors. As Lofgren *et al.* note in the context of influenza, "*the myriad theories accounting for seasonality (. . .) suggest that the elegant and predictable periodicity of nonpandemic influenza is caused by a less-than-straightforward interaction of many different factors,*" meaning that "*recognition of this complexity, as well as the likelihood that seasonality arises from many different factors, is essential for continued examination and elucidation of seasonality*" [17].

Given the severe methodological challenge of disentangling these interrelated factors, a more tractable solution is to approach seasonality holistically with the purpose of understanding its overall effects. In this study, we infer a single seasonality parameter, describing the amplitude of the yearly variation in the time-varying reproduction number, $R(t)$, for one climate region. While this single parameter does not disentangle the individual effects that

comprise a seasonal profile, it accounts for the overall magnitude of the seasonal effect on SARS-CoV-2 transmission and thereby provides valuable insights for long-term policy planning.

Since both COVID-19 incidence and the presence of governmental non-pharmaceutical interventions have waxed and waned in consecutive waves since early 2020, adjusting for NPI effects is crucial for any effort to infer the influence of seasonality on transmission, yet early analyses of environmental drivers have largely not done so [5, 7, 11].

Some studies of environmental factors have taken indirect approaches to avoiding the influence of NPIs on their environmental estimates. In a recent study of the association between humidity, temperature, and SARS-CoV-2 transmission in Europe and North America, Landier *et al.* exclude from their analysis any periods at least 28 days after the implementation of 'lockdown measures' [4]. Ma *et al.* do include periods where measures are in place, however, instead of directly utilising data on NPIs, they use smoothed spatial and temporal splines to indirectly adjust for their influence [3]. Smith *et al.* compare the role of temperature in the presence and absence of 'lockdown' in the United States but only include a binary measure of whether stay-at-home orders were in place [22].

By contrast, we directly adjust for the influence of specific interventions by extending two hierarchical Bayesian models of NPI effects from Brauner *et al.* [23] and Sharma *et al.* [24] to include a term representing the multiplicative seasonal influence on the effective reproduction number.

Employing a common technique in infectious disease modelling [25, 26], we assume the seasonal variation itself is described by a sinusoidal modulation. We re-analyse data from the aforementioned two studies [23, 24] while restricting the scope to European regions in the temperate climate zone, where we assume the seasonality effect to be comparable both in its environmental and behavioural causal components.

2 Methods

We build on two previously published, state-of-the-art NPI effectiveness models: Brauner *et al.* [23] and *Sharma et al.* [24]. To estimate the seasonal variation in transmission while adjusting for the effect of NPIs, we extend these "base models" to include a seasonality effect. We fit each of the two resulting "seasonal models" separately, on separate data (see below); we thus obtain two distinct estimates for the seasonality effect.

2.1 Models

2.1.1 Non-pharmaceutical intervention effectiveness models.   In this section, we give a short summary of the common core of both base models. Please refer to Brauner *et al.* [23] and *Sharma et al.* [24] for more detailed descriptions of the models and their differences. Additionally, see S2 Appendix for a schematic diagram of the model structures.

Both models use a data-driven, cross-region modelling approach, in which a Bayesian hierarchical model is fitted jointly to a large set of regions. The models estimate NPI effects by comparing the timing of interventions in each region to the subsequent numbers of cases or deaths. Because each region deployed various combinations of interventions in different orders and with different outcomes, this method can disentangle the effect of individual interventions.

More concretely, the models use case and death data from each region to "backward" infer the number of new infections at each point in time, which is then itself used to infer the reproduction numbers. The NPI effects are then estimated by relating the daily reproduction numbers to the active NPIs, across all days and regions. This relatively simple, data-driven
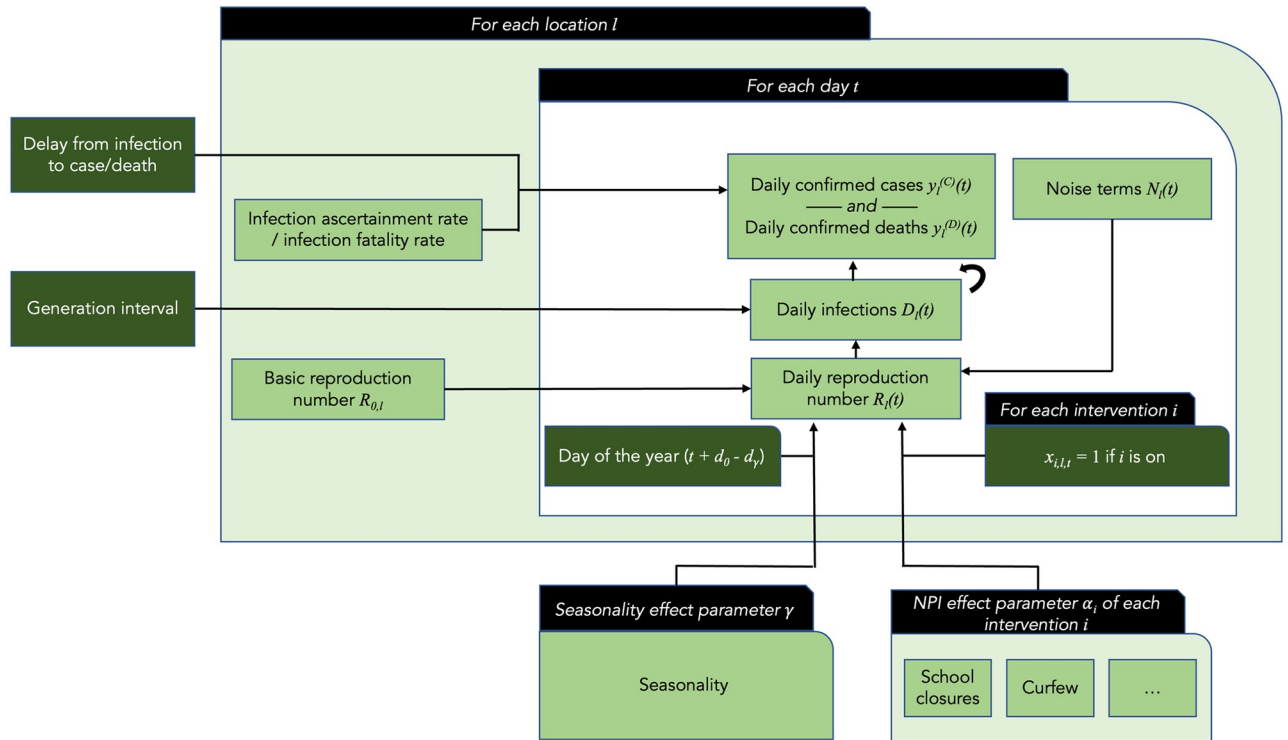
Fig 1. High-level overview of the common structure of both models used in this study. Dark nodes are observed, light nodes are inferred. See S2 Appendix for a detailed model graph for each individual model

approach makes it possible to sidestep the assumptions about contact patterns and intensity, infectiousness of different age groups, and so forth, which are typically required in modelling studies.

Fig 1 shows the basic model structure, including our adaptation to account for seasonality.

Both models assume the time-varying reproduction number $R(t)$ to be a product of R_0, the "natural" reproduction number when no NPIs are in place, multiple terms representing the effects of interventions, and noise terms modelling the influence of other, unobserved, factors on $R(t)$. The central model equation is thus:

$$\underbrace{R_l(t)}_{\text{Reproduction number in location } l \text{ at time } t} = \underbrace{R_{0,l}}_{R_0 \text{ in location } l} \overbrace{N_l(t)}^{R \text{ noise term}} \underbrace{\prod_{i=1}^{I} \exp(-\alpha_i\, x_{i,l,t})}_{\text{Effect due to active NPIs}} \quad (1)$$

where I is the number of NPIs, α_i is the "effect parameter" of intervention i, and $x_{i,l,t}$ are indicator variables, i.e. binary variables taking the value 1 if NPI i is active in location l at time t, and 0 otherwise. The reduction in $R(t)$ associated with each NPI (the "NPI effectiveness") can be computed as $1 - \exp(-\alpha_i)$. In Brauner *et al.*, but not Sharma *et al.*, the NPI effect parameter is allowed to vary slightly between different countries (partial pooling). The noise term $N_l(t)$ varies between the models; a log-normal multiplicative factor is used in Brauner *et al.* and a random walk-based multiplicative factor in Sharma *et al.* The noise term can be intuitively thought of as a random effect that accounts for residual variation in $R_l(t)$ not captured by the NPI effects.

The time-varying reproduction number $R_l(t)$ is then used to compute the change in the number of daily new infections $D_l(t)$. This relation between reproduction number and new infections is modelled differently in the two models. The Brauner *et al.* model uses a one-to-one correspondence between $R(t)$ and daily growth rates that holds early in a pandemic, and models separately the infections that will later become confirmed cases, and the infections that will become deaths. In contrast, the Sharma *et al.* model links $R_l(t)$ to $D_l(t)$ with a renewal process and models infections that will lead to cases/deaths jointly. The Sharma *et al.* model additionally includes additive noise on the number of infections, to account for the more fine-grained regions, and thus smaller regional infection numbers, in the data. In both models, the relation between $R_l(t)$ and $D_l(t)$ is moderated by the generation interval distribution (the generation interval is the time between successive infections in a chain of transmission).

New infections are then observed as confirmed cases or deaths, after some delay. For instance, the expected number of observed cases $\bar{y}_l^{(C)}(t)$ in location l at time t is computed as:

$$\bar{y}_l^{(C)}(t) = \sum_{\tau=0}^{31} D_l(\tau - t) \cdot P_C(\text{delay} = \tau) \cdot \text{IAR} \tag{2}$$

where P_C is the distribution over the delay between infection and case confirmation (truncated after 31 days because delays longer than this cutoff are very rare) and "IAR" is the infection ascertainment rate (i.e. the fraction of infected individuals that later test positive). The number of new COVID deaths is computed similarly, drawing on the delay between infection and death, and the infection fatality rate.

Finally, the observed cases $y_l^{(C)}(t)$ follow a negative binomial distribution with mean $\bar{y}_l^{(C)}(t)$ and an inferred dispersion parameter. This distribution reflects that small case numbers are more noisy and should therefore receive less weight. The number of observed deaths also follows a negative binomial distribution with a separate inferred dispersion parameter. Having separate dispersion parameters for cases and deaths ensures that they can be weighted differently if there is a difference in their output variance.

Both models require assumptions about priors for various variables, such as the NPI effectiveness or delay distributions. We use the same priors as in Brauner *et al.* [23] and *Sharma et al.* [24], where they have been justified (whenever possible) by reference to meta-analyses or patient line-list data.

2.1.2 Estimating seasonality effects. To account for seasonality, we substitute $R_l(t)$ with $R'_l(t)$ (adjusted for seasonality) and let each model infer a single seasonality amplitude parameter γ along with its other parameters. This minimal modification aims to preserve the demonstrated robustness of the original models [23, 24, 27]. We can thus estimate two separate seasonality effects; the final seasonality amplitude estimate is then pooled from the two models, equally sampling from their posterior distributions. Note the seasonal adjustment to $R_l(t)$ is *shared* across all locations l and therefore captures common dynamics between locations not explained by the location specific noise terms or NPIs.

We model seasonality as a sinusoidal multiplicative factor $\Gamma(t)$ to $R(t)$:

$$\Gamma(t) = 1 + \gamma \sin\left(2\pi \frac{t + d_0 - d_\gamma}{365} + \frac{\pi}{2}\right), \tag{3}$$

where γ is the intensity (amplitude) of the seasonal effect, d_γ is the day of the year of the highest seasonal effect on R, and d_0 is the first day of the respective dataset. Note that only γ is a random variable, but d_γ is assumed fixed (see below) and d_0 is a fixed property of the dataset. Our choice of the sinusoidal function for seasonality is motivated by an assumption that this function is a good fit for several of the upstream causal factors of seasonal forcing (S1 Appendix) in

temperate Europe, such as temperature and humidity. The sinusoidal function, and the equivalent cosine function, are commonly employed for mathematical modelling of seasonal transmission dynamics [28–30]. In S6 Appendix, section 1.6, we examine the sensitivity of our results to using Fourier series models of varying degrees as an alternative function for seasonal forcing.

We assume a single, common seasonal effect for countries in similar climates and relative proximity along dimensions such as income, political structure, and culture. While average temperatures clearly are different between countries within the region, with a strong dependence on latitude, the amplitude of the seasonal variation is assumed to be similar. For both models, the time and location-specific $R_l(t)$ is replaced with seasonal $R_l'(t)$:

$$R_l'(t) = R_l(t)\frac{\Gamma(t)}{\Gamma(0)}. \tag{4}$$

Note that we divide $\Gamma(t)$ by $\Gamma(0)$ to have $\Gamma(t)/\Gamma(0) = 1$ at $t = 0$ and $R_l'(0) = R_l(0)$, i.e. the seasonality multiplier is normalised to 1 at the start of the window of analysis. This implies that the priors over $\tilde{R}_l(0)$ need not be adjusted in either model. For both models, we assume an uniform prior $\gamma \sim U(-0.95, 0.95)$. (Sinusoidal seasonality is well-defined only for amplitudes $-1 \leq \gamma \leq 1$. We restrict γ to $-0.95 \leq \gamma \leq 0.95$ to ensure model numerical stability).

The amplitude of the cyclical seasonal variation (γ) can be converted to the reduction in transmission associated with going from the peak of winter to the peak of summer (i.e., peak-to-trough) as $R(\text{trough})/R(\text{peak}) = (1 - \gamma)/(1 + \gamma)$. Similarly, the amplitude can be directly converted to the expected reduction between adjacent seasons such as peak winter to mid-spring or mid-spring to peak summer (i.e., peak-to-mid).

Our analysis utilises January 1 as the seasonal peak day d_γ, as this date is both close to the center of a stable range of d_γ in the sensitivity analysis of S6 Appendix, as well as close to January 3, the median peak date inferred by a model with variable d_γ in S6 Appendix. Note that while we show January 1 to be a robust choice of d_γ, we are not aiming at determining its exact value.

2.1.3 Inference.   We infer the unobserved variables in our models using the No-U-Turn Sampler (NUTS) [31], a standard Markov chain Monte Carlo sampling algorithm; we use 250 tuning samples per chain. As explained in Eqs 3 and 4, we add γ as a fully-pooled (global) hidden variable to each of the two models and inferred in the same way as other variables. We fit each model separately, on different data (see below). For each model we obtain 5000 posterior samples in 4 independent chains (1250 in each chain). The Gelman-Rubin's $\hat{R}$ statistic across the 4 chains was under 1.02 for all variables in both models, with no divergences.

2.3 Data

We fit each model separately, on the data and time period for which the model was originally created. Brauner *et al.* includes data on the implementation and lifting of several NPIs in 41 countries between 22nd January and 30th May 2020. We restrict the dataset of Brauner *et al.* to the 29 countries in the temperate European region (see S2 Appendix). Sharma *et al.* contains data on 17 NPIs in 114 subnational regions in 7 European countries (Austria, Czech Republic, England, Germany, Italy, Netherlands, Switzerland), and covers the period from 1st August 2020 to 9th January 2021. We use the Sharma *et al.* dataset without any modifications, for a total of 143 regions of analysis. We follow the same pre-processing steps as in the original datasets. In particular, for Brauner *et al*, we mask case numbers before a country had reached 100 confirmed cases and fatality numbers before a country had reached 10 deaths, to prevent bias from case/death importation (this is not necessary for the Sharma *et al.* model due to its noise
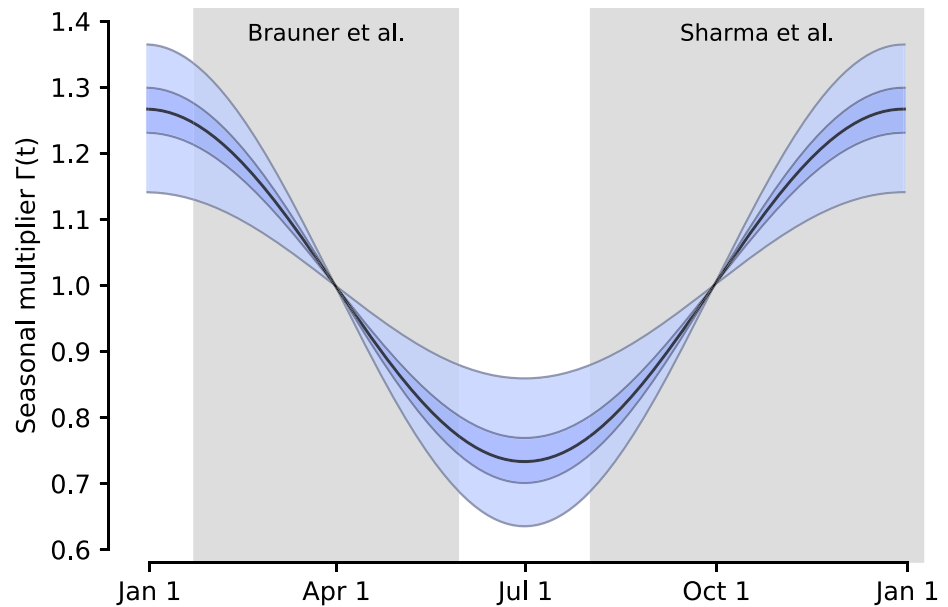
Fig 2. The inferred seasonal *R* multiplier Γ(*t*) of the combined models estimate, with 50% and 95% credible intervals. Gray boxes indicate data range of each dataset, i.e. 22nd January to 30th May 2020 for Brauner *et al.* and 1st August 2020 to 9th January 2021 for Sharma *et al.* The zero-width credible intervals around April 1 and October 1 owe to the fact that we model Γ(*t*) as a seasonal multiplier for *R* relative to the mid-spring and mid-fall, respectively, which implies that Γ(*t*) is assumed to be exactly equal to one for these dates.

on infections). For Sharma *et al.*, we exclude any data points where the prevalence of variants of concern exceeded 10% during a given day, to mitigate potential bias introduced by more transmissible strains.

Data on confirmed cases and deaths were taken from the Johns Hopkins CSSE COVID-19 Dataset [32] and government websites (see Supplementary Table S4 in [24]).

3 Results

Using two model structures and datasets on non-pharmaceutical interventions covering 72% of the 2020–2021 period in Europe, we estimate the seasonality parameter γ and the time-varying seasonal multiplier Γ(*t*) (Fig 2). See S3 Appendix for details.

Our combined estimates from the two models are consistent with a reduction in *R* of 24.7% to 53.4% (95% CI) from January 1, the peak of winter to July 1, the peak of summer, with a median reduction of 42.1% (Fig 3 and detailed results in S3 Appendix). The combined estimate distribution is an equal-weight mixture of the posterior distributions of the two models.

Modelling seasonality alongside non-pharmaceutical interventions allows us to gain a sense of the epidemiological importance of environmental factors. We find that the transition from winter to summer is associated with a reduction in transmission that is comparable to or greater than the effects of individual interventions, but less than the total effect of combined interventions (Fig 4).

Fig 4 compares our seasonality estimates to the original effect estimates from Brauner *et al.*, as their robustness is well-established [23, 27]. Although these estimates are based on analysis that included countries outside temperate Europe, we find that restricting our analysis to temperate regions has little effect on the inferred total effect of NPIs and thus should not invalidate the comparison (S6 Appendix, section 1.4).
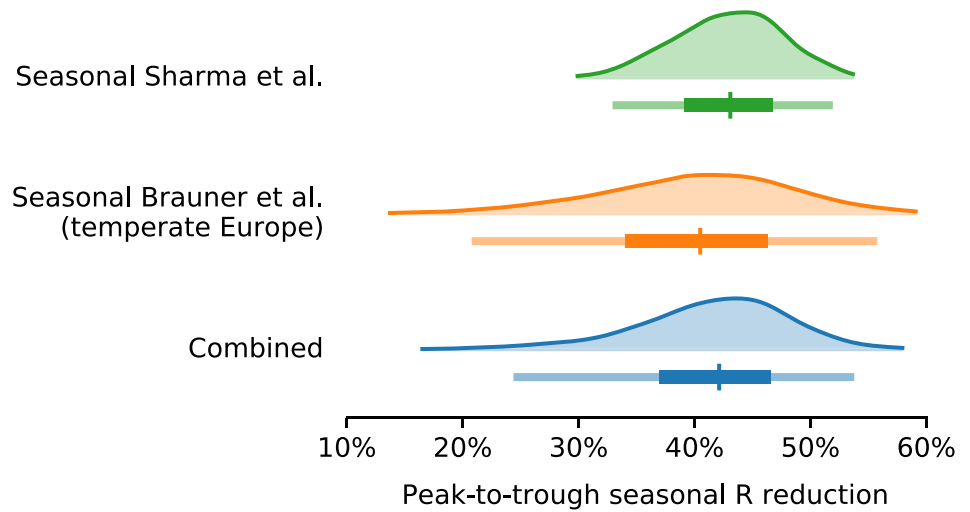
Fig 3. Posterior distributions of the *R* reduction on July 1 relative to January 1 with median, 50% and 95% credible intervals.

Beyond NPIs, voluntary changes in behaviour and contact patterns constitute important influences on the reproduction rate. As noted, seasonal variation in behavioural patterns such as time spent indoors are an important component of our holistic conception of seasonality (S1 Fig; S1 Appendix). However, if there are behavioural changes over time that are causally
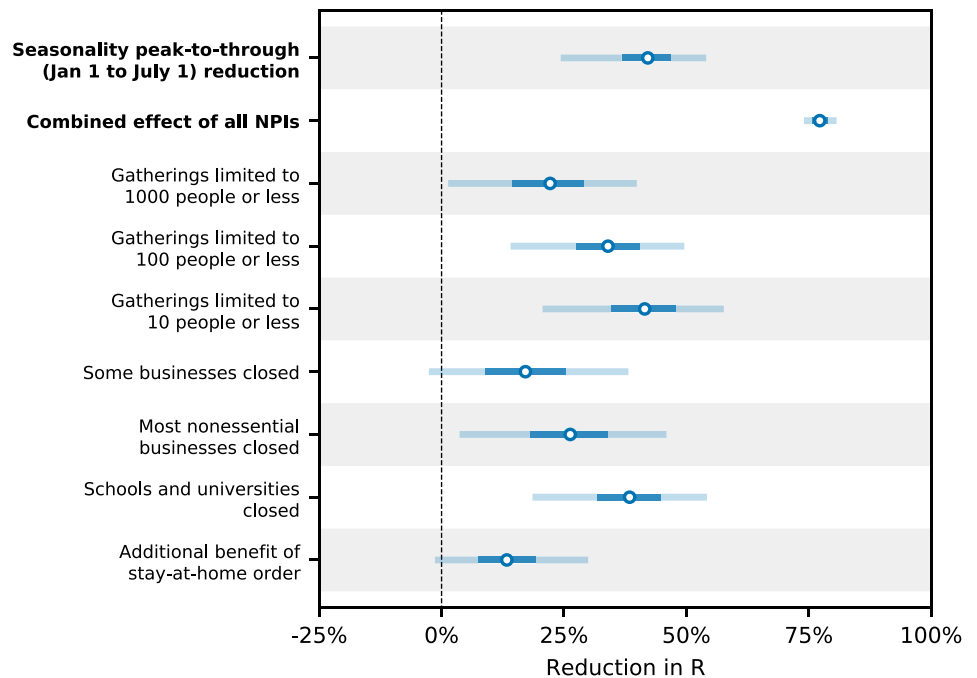
Fig 4. Comparison of the inferred peak-to-trough *R* reduction effect of seasonality (combined from both models) to the NPI reductions inferred by Brauner et al. [23], with 50% and 95% CIs. The seasonal effect is lower than the combined NPI effect but higher than or comparable to the individual NPI effects.

unrelated with the transition between seasons, these may be mistakenly attributed to a causal effect of seasonal forcing. If, for example, voluntary personal protective behaviours beyond compliance with NPIs increased during the spring and decreased during the autumn, this would provide an alternative explanation of the respective fall and rise in transmission during those seasons. To examine whether behavioural changes over the course of the pandemic influenced our results, we incorporated country-level mobility scores into the model as an additional independent variable (see S4 Appendix). We find that our seasonality estimates are robust to adjusting for mobility trends. Specifically, our estimate of γ changes by less than 1.64% with the adjustment for mobility. While this finding strongly suggests that the observed seasonal pattern cannot be explained by unrelated changes in behaviour, it should be noted that data on mobility may not capture every relevant aspect of behavioural trends and that we therefore cannot conclusively rule out the possibility that our estimates are influenced by changes in behaviour.

Another important consideration is that the effect of seasonality on transmission can be mediated through many causal pathways (see S1 Fig), such as the interaction between seasonality and active interventions. For example, colder seasons may lead to increased transmission via more gatherings occurring indoors relative to outdoors. Notably, this phenomenon would be less pronounced when gatherings are banned, leading to a potential interaction between seasonality and NPI implementation. Our results provide the *average* seasonality effect across European countries in the first year of the pandemic. If countries implement very different NPIs (or none), the seasonality effect may be different as well. S6 Appendix, section 1.5, examines how our results are affected by including a term for the interaction between seasonality and NPIs; we find that neither our estimates for seasonality nor for NPIs are very sensitive to the inclusion of this interaction.

Incorporating seasonality into models of NPI effectiveness may also improve their estimates by explaining residual variation in the inferred reproduction rate. A key advance in the model proposed by Sharma *et al.* was the incorporation of a stochastic random walk process on the basic reproduction number to flexibly account for trends in transmission due to unobserved factors [24]. We find that including the seasonality term reduces the magnitude and asymmetry of the random walk considerably, thereby reducing the internal model variation (S5 Appendix). Specifically, we find that the mean square displacement (MSD) of the random walk in log-space is 0.131 for the non-seasonal model and 0.072 for the seasonal model. These results suggest a considerable amount of the residual variation can be explained by a common seasonality profile. S5 Appendix compares the NPI effects from Brauner *et al.* and Sharma *et al.* with and without the inclusion of seasonality in the model.

Estimates of seasonality and NPI effects are sensitive to modelling choices [23, 24, 27]. It is therefore vital to include a sensitivity analyses of free parameters and inputs to ensure consistent results. Relying on the demonstrated robustness of the original models, we focus primarily on the parameter that we introduce in the form of peak seasonality day. We find that the inferred mean peak-to-trough reduction in R varies by less than 5% across all the analysed peak seasonality dates in December and January (S6 Appendix, section 1.2). Although the seasonality magnitude is somewhat sensitive to setting the winter peak to different dates in February, these dates are considerably later in the year than the median peak date inferred in our sensitivity analysis, January 3 (see S6 Appendix, section 1.1).

Since the seasonality term we introduce is directly related to $R_l(t)$ through Eq (4), we also examine the sensitivity of our results the mean initial R_0 prior. We find that our results are robust to univariate variation in this parameter, with the seasonal Sharma *et al.* model being the most sensitive (S6 Appendix, section 1.3).

4 Discussion

In this study, we sidestepped the intractable methodological challenge of evaluating various highly interrelated seasonal factors, and instead provide a precise estimate for the overall seasonal variation in SARS-CoV-2 transmission in temperate Europe, while adjusting for both non-pharmaceutical interventions and overall changes in mobility patterns. The strong associations we observe match the clear seasonal patterns of other respiratory viruses [2]. While reductions in reproduction rates and case numbers are not directly comparable, another recent analysis by Chen *et al.* [7] infers a 64% reduction in cases from one season to the next based on a cross-sectional regression at a single point in time, similarly suggesting a significant role of environmental factors. The general magnitude of our results is also in line with previous assumptions about the magnitude of SARS-CoV-2 seasonal forcing. For example, Kissler *et al.* assume that the reduction of SARS-CoV-2 R_0 between winter and summer peaks ranges from 10% to 40% [1], while Neher *et al.* [33] assume values of γ between 0.3 and 0.7. Moreover, recent analyses have suggested a role of environmental factors in the B.1.1.7 lineage transmission intensity and that such factors may differentially affect the transmission of different variants of concern [6].

It is important to note that our results are not inconsistent with widespread outbreaks in warmer regions, nor do they imply that temperate regions cannot face surges in transmissions during summer periods. Despite moderate seasonal forcing, the time-varying reproduction number can remain well above 1 during the peak of the summer, particularly given high incidence of more transmissible variants such as lineage B.1.617.2 [34]. Indeed, in certain parts of Europe, R remained above 1 even during the warmer periods of the study window and transmission intensity currently remains high in several warmer regions across the world. Previous modelling has suggested that population immunity limits the role of environmental factors [10]. Consequently, vaccination rates, non-pharmaceutical interventions, and the prevalence of more transmissible variants will continue to be important determinants of transmission throughout the year.

Moreover, this study utilised variation in environmental and behavioural factors across time while holding the climate zone constant, and the observed results may not directly translate to comparisons across regions holding the season constant. In other words, the relationship between cooler periods and transmission within the temperate zone does not necessarily imply an exactly similar association between regional climate and transmission rates at any given point in time. This is because latitude is correlated with a wide range of epidemiological, demographic, and societal factors, each of which may affect transmission.

A major limitation of our analysis is that it relies on data from only one complete period of seasonality. We present the inferred seasonality estimates as the best estimate given the available data. Moreover, since our analysis focused exclusively on European regions in the temperate climate zone, the findings may not generalise to other climates, particularly as we have not identified the relative contributions of different causal mechanisms. Other respiratory infections show less seasonality in tropical regions relative to temperate regions as well as seasonal patterns with different peak timings, for example, during the monsoon season [2, 35]. Further research can shed light on the extent to which this is the case for SARS-CoV-2, and on the interaction between seasonality and latitude within climate regions. Moreover, further research can shed light on whether and how changes in population immunity outside our window of analysis may interact with seasonal effects.

More generally, this observational study demands caution when drawing conclusions about causality. Our analysis did not attempt to disentangle the various plausible causal pathways through which seasonality may affect transmission, and both environmental and behavioural

factors can vary over the years. For example, behavioural patterns throughout the first year of the pandemic were likely exceptional, and while some behavioural changes are closely tied to modelled NPIs and thus do not bias our analysis, other relevant behavioural aspects may differ in subsequent years. Consequently, a granular focus on specific factors such as temperature, humidity, and behaviour is required for short-term prediction to inform policy.

Notwithstanding these limitations, a parsimonious form for seasonality may be adequate to understand variations over time and aid long-term policy planning. Even without disentangling the underlying factors, our approach to incorporating seasonality can augment modelling efforts to anticipate changes in transmission patterns more reliably, particularly when adjusting for important factors such as non-pharmaceutical interventions.

For such forward-looking analyses of SARS-CoV-2 seasonality, it should be noted that our inferred seasonal associations do not include two factors that play significant roles in the seasonality of other respiratory viruses. First, we treat school closures, including for holidays, as NPIs in our model due to the role of closing educational institutions in the epidemic responses of many countries. This means that any effects of closing schools are attributed to the school NPI, rather than to seasonality. This is noteworthy considering that school calendars are considered an important driver of seasonality for other respiratory viruses [19, 36]. Consequently, the full extent of seasonality would likely be greater if it is construed to include school calendars. Second, the seasonal variation of some respiratory viruses, such as influenza, owes to a combination of both the direct seasonal forcing from biological and behavioural factors as well as the indirect influence of waning population immunity [37]. Given what is known about the robustness of acquired immunity within the first year of SARS-CoV-2 infection [38], the patterns we observe likely owe almost entirely to seasonal forcing. Going forward, the long-term seasonality of SARS-CoV-2 will depend in part on developments in population immunity as well as on the emergence of variants.

Failing to account for seasonality may lead to a Panglossian outlook causing grave policy errors. For instance, a reduction in transmission over the summer could be misinterpreted as the result of population immunity alone [39], rather than the more plausible combination of immunity, seasonality, and other factors. Such an interpretation may lead to diminished support for continued vaccination efforts and booster vaccine uptake, or the premature discontinuation of non-pharmaceutical interventions. Each of these outcomes could ultimately result in an inadequate preparation for a resurgence during the colder months. At the same time, overestimating the role of environmental factors may be equally perilous. If policymakers anticipate a greater reduction in the summer months due to seasonality than will actually occur, such an interpretation could similarly lead to inadequate policy- and behavioural measures. We hope that our approach to modelling seasonality will allow for better calibrated policy responses for seasonal endemic and epidemic pathogens alike.

Supporting information

S1 Fig. Diagram of potential causal pathways for SARS-CoV-2 seasonality. A complex web of environmental, biological, and behavioural factors contribute to the seasonality of respiratory viruses. Note that this diagram excludes school calendars, as these are subsumed under non-pharmaceutical interventions for the purposes of our analysis.
(PDF)

S1 Appendix. Causal pathways for SARS-COV-2 seasonality.
(PDF)

S2 Appendix. Datasets, model overview, and implementation.
(PDF)

S3 Appendix. Detailed results.
(PDF)

S4 Appendix. Sensitivity to adjustment for mobility.
(PDF)

S5 Appendix. Incorporating seasonality for NPI effect estimation.
(PDF)

S6 Appendix. Sensitivity analysis.
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Tomáš Gavenčiak, Jan Kulveit.

Data curation: Mrinank Sharma, Sören Mindermann, Jan Brauner.

Formal analysis: Tomáš Gavenčiak, Gavin Leech, Jan Kulveit.

Methodology: Tomáš Gavenčiak, Gavin Leech, Mrinank Sharma, Sören Mindermann, Samir Bhatt, Jan Brauner, Jan Kulveit.

Project administration: Tomáš Gavenčiak, Joshua Teperowski Monrad, Jan Kulveit.

Supervision: Tomáš Gavenčiak, Samir Bhatt, Jan Brauner, Jan Kulveit.

Validation: Tomáš Gavenčiak, Gavin Leech.

Visualization: Tomáš Gavenčiak, Gavin Leech.

Writing – original draft: Tomáš Gavenčiak, Joshua Teperowski Monrad, Gavin Leech, Jan Kulveit.

Writing – review & editing: Tomáš Gavenčiak, Joshua Teperowski Monrad, Gavin Leech, Mrinank Sharma, Sören Mindermann, Samir Bhatt, Jan Brauner, Jan Kulveit.

References

1. Kissler SM, Tedijanto C, Goldstein E, Grad YH, Lipsitch M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. Science. 2020; 368(6493):860–868. https://doi.org/10.1126/science.abb5793 PMID: 32291278

2. Moriyama M, Hugentobler WJ, Iwasaki A. Seasonality of respiratory viral infections. Annual review of virology. 2020; 7:83–101. https://doi.org/10.1146/annurev-virology-012420-022445 PMID: 32196426

3. Ma Y, Pei S, Shaman J, Dubrow R, Chen K. Role of meteorological factors in the transmission of SARS-CoV-2 in the United States. Nat Commun. 2021; 12(3602):1–9. https://doi.org/10.1038/s41467-021-23866-7 PMID: 34127665

4. Landier J, Paireau J, Rebaudet S, Legendre E, Lehot L, Fontanet A, et al. Cold and dry winter conditions are associated with greater SARS-CoV-2 transmission at regional level in western countries during the first epidemic wave. Sci Rep. 2021; 11(12756):1–13. https://doi.org/10.1038/s41598-021-91798-9 PMID: 34140557

5. Sajadi MM, Habibzadeh P, Vintzileos A, Shokouhi S, Miralles-Wilhelm F, Amoroso A. Temperature, humidity, and latitude analysis to estimate potential spread and seasonality of coronavirus disease 2019 (COVID-19). JAMA network open. 2020; 3(6):e2011834–e2011834. https://doi.org/10.1001/jamanetworkopen.2020.11834 PMID: 32525550

6. Smith TP, Dorigatti I, Mishra S, Volz E, Walker PGT, Ragonnet-Cronin M, et al. Environmental drivers of SARS-CoV-2 lineage B.1.1.7 transmission intensity. medRxiv. 2021.

7. Chen S, Prettner K, Kuhn M, Geldsetzer P, Wang C, Bärnighausen T, et al. Climate and the spread of COVID-19. Scientific Reports. 2021; 11(1):1–6. https://doi.org/10.1038/s41598-021-87692-z PMID: 33907202

8. Poirier C, Luo W, Majumder MS, Liu D, Mandl KD, Mooring TA, et al. The role of environmental factors on transmission rates of the COVID-19 outbreak: an initial assessment in two spatial scales. Scientific reports. 2020; 10(1):1–11. https://doi.org/10.1038/s41598-020-74089-7

9. Carlson CJ, Gomez AC, Bansal S, Ryan SJ. Misconceptions about weather and seasonality must not misguide COVID-19 response. Nature Communications. 2020; 11(1):1–4. https://doi.org/10.1038/s41467-020-18150-z

10. Baker RE, Yang W, Vecchi GA, Metcalf CJE, Grenfell BT. Susceptible supply limits the role of climate in the early SARS-CoV-2 pandemic. Science. 2020; 369(6501):315–319. https://doi.org/10.1126/science.abc2535 PMID: 32423996

11. McClymont H, Hu W. Weather Variability and COVID-19 Transmission: A Review of Recent Research. International journal of environmental research and public health. 2021; 18(2):396. https://doi.org/10.3390/ijerph18020396 PMID: 33419216

12. Shaman J, Pitzer VE, Viboud C, Grenfell BT, Lipsitch M. Absolute humidity and the seasonal onset of influenza in the continental United States. PLoS Biol. 2010; 8(2):e1000316. https://doi.org/10.1371/journal.pbio.1000316 PMID: 20186267

13. Yang W, Elankumaran S, Marr LC. Relationship between humidity and influenza A viability in droplets and implications for influenza's seasonality. PloS one. 2012; 7(10):e46789. https://doi.org/10.1371/journal.pone.0046789 PMID: 23056454

14. Lowen AC, Mubareka S, Steel J, Palese P. Influenza virus transmission is dependent on relative humidity and temperature. PLoS Pathog. 2007; 3(10):e151. https://doi.org/10.1371/journal.ppat.0030151 PMID: 17953482

15. Kudo E, Song E, Yockey LJ, Rakib T, Wong PW, Homer RJ, et al. Low ambient humidity impairs barrier function and innate resistance against influenza infection. Proceedings of the National Academy of Sciences. 2019; 116(22):10905–10910. https://doi.org/10.1073/pnas.1902840116 PMID: 31085641

16. Foxman EF, Storer JA, Vanaja K, Levchenko A, Iwasaki A. Two interferon-independent double-stranded RNA-induced host defense strategies suppress the common cold virus at warm temperature. Proceedings of the National Academy of Sciences. 2016; 113(30):8496–8501. https://doi.org/10.1073/pnas.1601942113 PMID: 27402752

17. Lofgren E, Fefferman NH, Naumov YN, Gorski J, Naumova EN. Influenza seasonality: underlying causes and modeling theories. Journal of virology. 2007; 81(11):5429–5436. https://doi.org/10.1128/JVI.01680-06 PMID: 17182688

18. Willem L, Van Kerckhove K, Chao DL, Hens N, Beutels P. A nice day for an infection? Weather conditions and social contact patterns relevant to influenza transmission. PloS one. 2012; 7(11):e48695. https://doi.org/10.1371/journal.pone.0048695 PMID: 23155399

19. Ewing A, Lee EC, Viboud C, Bansal S. Contact, travel, and transmission: The impact of winter holidays on influenza dynamics in the United States. The Journal of infectious diseases. 2017; 215(5):732–739. https://doi.org/10.1093/infdis/jiw642 PMID: 28031259

20. Eames KT, Tilston NL, Brooks-Pollock E, Edmunds WJ. Measured dynamic social contact patterns explain the spread of H1N1v influenza. PLoS Comput Biol. 2012; 8(3):e1002425. https://doi.org/10.1371/journal.pcbi.1002425 PMID: 22412366

21. Lipsitch M, Viboud C. Influenza seasonality: lifting the fog. Proceedings of the National Academy of Sciences. 2009; 106(10):3645–3646. https://doi.org/10.1073/pnas.0900933106 PMID: 19276125

22. Smith TP, Flaxman S, Gallinat AS, Kinosian SP, Stemkovski M, Unwin HJT, et al. Temperature and population density influence SARS-CoV-2 transmission in the absence of nonpharmaceutical interventions. Proc Natl Acad Sci USA. 2021; 118(25):e2019284118. https://doi.org/10.1073/pnas.2019284118 PMID: 34103391

23. Brauner JM, Mindermann S, Sharma M, Johnston D, Salvatier J, Gavenčiak T, et al. Inferring the effectiveness of government interventions against COVID-19. Science. 2021; 371 (6531). https://doi.org/10.1126/science.abd9338 PMID: 33323424

24. Sharma M, Mindermann S, Rogers-Smith C, Leech G, Snodin B, Ahuja J, et al. Understanding the effectiveness of government interventions against the resurgence of COVID-19 in Europe. Nature communications. 2021; 12(1):1–13. https://doi.org/10.1038/s41467-021-26013-4 PMID: 34611158

25. Stolwijk A, Straatman H, Zielhuis G. Studying seasonality by using sine and cosine functions in regression analysis. Journal of Epidemiology & Community Health. 1999; 53(4):235–238. https://doi.org/10.1136/jech.53.4.235 PMID: 10396550

26. Allen LJ, Brauer F, Van den Driessche P, Wu J. Mathematical epidemiology. vol. 1945. Springer; 2008.

27. Sharma M, Mindermann S, Brauner J, Leech G, Stephenson A, Gavenčiak T, et al. How Robust are the Estimated Effects of Nonpharmaceutical Interventions against COVID-19? In: Advances in Neural Information Processing Systems. vol. 33; 2020. p. 12175–12186. Available from: https://proceedings.neurips.cc/paper/2020/file/8e3308c853e47411c761429193511819-Paper.pdf.

28. Jódar L, Villanueva RJ, Arenas A. Modeling the spread of seasonal epidemiological diseases: theory and applications. Mathematical and Computer Modelling. 2008; 48(3-4):548–557. https://doi.org/10.1016/j.mcm.2007.08.017

29. White L, Mandl J, Gomes M, Bodley-Tickell A, Cane P, Perez-Brena P, et al. Understanding the transmission dynamics of respiratory syncytial virus using multiple time series and nested models. Mathematical biosciences. 2007; 209(1):222–239. https://doi.org/10.1016/j.mbs.2006.08.018 PMID: 17335858

30. Weber A, Weber M, Milligan P. Modeling epidemics caused by respiratory syncytial virus (RSV). Mathematical biosciences. 2001; 172(2):95–113. https://doi.org/10.1016/S0025-5564(01)00066-9 PMID: 11520501

31. Hoffman MD, Gelman A, et al. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. J Mach Learn Res. 2014; 15(1):1593–1623.

32. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. The Lancet infectious diseases. 2020; 20(5):533–534. https://doi.org/10.1016/S1473-3099(20)30120-1 PMID: 32087114

33. Neher RA, Dyrdak R, Druelle V, Hodcroft EB, Albert J. Potential impact of seasonal forcing on a SARS-CoV-2 pandemic. Swiss medical weekly. 2020; 150 (1112). PMID: 32176808

34. Mishra S, Mindermann S, Sharma M, Whittaker C, Mellan TA, Wilton T, et al. Changing composition of SARS-CoV-2 lineages and rise of Delta variant in England. EClinicalMedicine. 2021; 39:101064. https://doi.org/10.1016/j.eclinm.2021.101064 PMID: 34401689

35. Price RHM, Graham C, Ramalingam S. Association between viral seasonality and meteorological factors. Scientific Reports. 2019; 9(1). https://doi.org/10.1038/s41598-018-37481-y PMID: 30700747

36. Chao DL, Halloran ME, Longini IM. School opening dates predict pandemic influenza A (H1N1) outbreaks in the United States. The Journal of infectious diseases. 2010; 202(6):877–880. https://doi.org/10.1086/655810 PMID: 20704486

37. Grenfell BT, Pybus OG, Gog JR, Wood JL, Daly JM, Mumford JA, et al. Unifying the epidemiological and evolutionary dynamics of pathogens. science. 2004; 303(5656):327–332. https://doi.org/10.1126/science.1090727 PMID: 14726583

38. Sette A, Crotty S. Adaptive immunity to SARS-CoV-2 and COVID-19. Cell. 2021;. https://doi.org/10.1016/j.cell.2021.01.007 PMID: 33497610

39. Okell LC, Verity R, Watson OJ, Mishra S, Walker P, Whittaker C, et al. Have deaths from COVID-19 in Europe plateaued due to herd immunity? The Lancet. 2020; 395(10241):e110–e111. https://doi.org/10.1016/S0140-6736(20)31357-X PMID: 32534627