scientific reports

OPEN

Differences in COVID-19 cyclicity and predictability among U.S. counties and states reflect the effectiveness of protective measures

Claudio Bozzuto[1,3]✉ & Anthony R. Ives[2,3]

During the COVID-19 pandemic, many quantitative approaches were employed to predict the course of disease spread. However, forecasting faces the challenge of inherently unpredictable spread dynamics, setting a limit to the accuracy of all models. Here, we analyze COVID-19 data from the USA to explain variation among jurisdictions in disease spread predictability (that is, the extent to which predictions are possible), using a combination of statistical and simulation models. We show that for half the counties and states the spread rate of COVID-19, $r(t)$, was predictable at most 9 weeks and 8 weeks ahead, respectively, corresponding to at most 40% and 35% of an average cycle length of 23 weeks and 26 weeks. High predictability was associated with high cyclicity of $r(t)$ and negatively associated with R_0 values from the pandemic's onset. Our statistical evidence suggests the following explanation: jurisdictions with a severe initial outbreak, and where individuals and authorities took strong and sustained protective measures against COVID-19, successfully curbed subsequent waves of disease spread, but at the same time unintentionally decreased its predictability. Decreased predictability of disease spread should be viewed as a by-product of positive and sustained steps that people take to protect themselves and others.

Human societies have always experienced outbreaks of infectious diseases, and disease epidemics are expected to emerge or re-emerge more frequently in the future[1–3]. The COVID-19 pandemic, caused by the SARS-CoV-2 virus, showed the limited strategies and actions humans have at their disposal to prevent outbreaks of emerging diseases, and the suffering and death once a disease starts spreading[2,4].

If a disease outbreak cannot be prevented, public health officials and politicians will try to swiftly implement measures to help minimize disease-related suffering and death[5,6]. Such measures can range from preparing and re-organizing medical infrastructure (e.g., increasing personnel for intensive care units) to enacting non-pharmaceutical interventions (NPIs), either as mandates or as recommendations to the public. For impending or unfolding disease outbreaks, forecasts have proven helpful for emergency planning[6,7]. To match the time required to plan and implement mitigation actions for public health needs, however, the lead-time of the forecasts typically ranges from one week to two or more months[6,8]. Long-term forecasts are important to prepare for resurgences of the disease, as has happened worldwide with COVID-19[9,10], and also to justify severe NPI mandates such as lockdowns: mandates that disrupt social and economic systems can be justified if the course of the disease spread is expected to last months and lead to a high death toll. For re-emerging influenza outbreaks, Viboud and Vespignani (Ref.[8], p. 2804) aptly use a weather forecast analogy: "the influenza forecasting community will need to offer weather forecasts as well as climate predictions".

The COVID-19 pandemic has spurred an unprecedented effort to quantitatively understand disease spread and forecast spread dynamics to help public health officials implement protective measures such as NPIs (Ref.[11], and references therein). Nonetheless, these efforts face the challenge that the predictability of COVID-19 spread may be inherently limited. Here, we use the definition that "predictability is the study of the extent to which events can be predicted" (Ref.[12], p. 2425). Several epidemiological studies have addressed the fundamental limit to

[1]Wildlife Analysis GmbH, Oetlisbergstrasse 38, 8053 Zurich, Switzerland. [2]Department of Integrative Biology, University of Wisconsin-Madison, Madison, WI 53706, USA. [3]These authors contributed equally: Claudio Bozzuto and Anthony R. Ives. ✉email: bozzuto@wildlifeanalysis.ch

predictability of disease spread using model-free, entropy-based approaches (e.g. Ref.[13,14]). For example, Scarpino and Petri[14] found that for nine human diseases, there is a barrier to predictability, but that single outbreaks are in general predictable and that predictability depends in part on the basic reproduction number, R_0. Furthermore, these authors found considerable variation in predictability among jurisdictions for single diseases. In comparison, assessments of realized predictability (i.e. forecast accuracy) for influenza and COVID-19 outbreaks have shown that four weeks seems to be the forecast horizon beyond which the dynamics are hard to predict[8,15–17], implying that predicting COVID-19 resurgences two months in advance may be futile.

Model-free approaches address predictability with methods heavily relying on information theory. We worry that public health officials facing an epidemic and planning for public health responses need more concrete assessments of the limits to predictability as well as the factors that might determine this predictability. Here, we use time series models to statistically fit disease spread dynamics, and then analyze the predictability of the fitted models using the measure predictive power, $PP(t)$, rooted in information theory and developed in climatology[18] (see also Ref.[12]). An advantage of our approach is that we can associate predictability to specific dynamical patterns observed during the pandemic, like cyclic dynamics, which potentially lead to more accurate predictions (e.g. Ref.[19]).

For centuries it has been known that infectious disease outbreaks resurge regularly over time (e.g. Ref.[20]). Resurgent outbreaks can have many causes such as seasonality, school terms, or new pathogen variants (Ref.[20–22], and references therein). For COVID-19, too, the dynamics are characterized by 'waves' or cycles, not only in the USA but throughout the world, and different cyclic patterns have been documented, for example, at weekly and seasonal time scales[9,10,19]. Moreover, for many countries in both hemispheres additional cycles occur with a period of approximately 4 months (3–6 months), similar to other communicable (viral) diseases like the Spanish flu from 1918 (approximately 5 months; Ref.[10]). Mitchell and Zhang[10] speculate that these cycles are caused by virus-host feedbacks, and other studies show that models incorporating behavioral responses to limit disease spread can show cyclic dynamics when these responses occur with a time delay[23–25]. We investigate the cyclic dynamics of COVID-19 using a stochastic epidemiological model to understand how human responses to infection rates may affect cyclicity and predictability of disease spread.

Our overall goal is to understand the high variation among counties and states in predictability of COVID-19 spread dynamics during the period after its establishment (May 2020) and before vaccinations became widely available (February 2021). We use this variation to develop an explanation for cyclicity and predictability of the COVID-19 pandemic.

Materials and methods

Estimation of COVID-19 spread rate $r(t)$.
We base our analyses on the disease spread rate, $r(t)$, of COVID-19 in the USA, estimated at the county and state levels (henceforth jurisdictions) using weekly death counts[26] from 9 May 2020 to 12 February 2021 (40 weeks). We did not consider the initial outbreak (March-early May 2020) because there was pronounced among-jurisdiction variation in the time of onset[27], and because protective measures (individual behavior and NPIs) built up differently during the first outbreak[28]. We ended the data on 12 February 2021 because vaccinations had started to influence the disease transmission and death rates[29]. Our estimates of $r(t)$ depend on the weekly difference between two adjacent log-transformed death counts; thus, at the original scale death count $D(t) \propto D(t-1)\exp(r(t-1))$. We used death counts rather than reported cases of disease because death data are less likely to give biased estimates of spread rates than case data[30]. Furthermore, predicting death rates is critical for health care in terms of both direct human costs and medical preparedness for increases in critical cases of infection. At the state level, we used data for the 49 conterminous states in the USA (including the District of Columbia), while at the county level we selected from these states the 100 counties with the highest population size to maximize estimation accuracy.

To estimate $r(t)$ using the entire time series, we used a previously published time-varying autoregressive model in state-space form[27]; we present a summary, including model equations, in the Supplementary Information, section *Estimation of COVID-19 spread rate $r(t)$*. Briefly, the model estimates the unobserved difference between adjacent log-transformed observed death counts. These differences constitute the time-varying spread rate, resulting in jurisdiction-specific time series to be analyzed further (see below). This type of approach of reconstructing the spread rate is not often used in epidemiological studies, but it has the advantage of being statistically robust even when the data (death counts) are few, and it makes the minimum number of assumptions that could affect the estimates in unexpected ways[27,31]. An additional advantage of using a state-space model (fitted using the Kalman filter[32]) is the explicit inclusion of measurement error in the observed death rates; this is important for jurisdictions with low death tolls. Finally, we used the Kalman smoother[32] to produce the final $r(t)$ time series. The Kalman filter gives the maximum likelihood parameter estimates for the time series model of $r(t)$, while the Kalman smoother uses these estimates plus all of the data available in the time series to obtain the best estimates of $r(t)$ at each time point; thus, the Kalman smoother seizes all information available after fitting and 'retrospectively adjusts' the values of $r(t)$ in their entirety[32]. Figure 1 shows example data and estimated $r(t)$ time series of three counties, and the Supplementary Figs. S1–S2 show all estimated $r(t)$ time series at the county and state levels grouped by similarity of the spread dynamics. These fits of $r(t)$ are the best 20:20 hindsight estimates that use all data in the time series. For real-time forecasting, short time series will cause uncertainty in model parameter estimates and hence $r(t)$, but because we are interested in the inherent limit to predictability of the process underlying $r(t)$, we use the best possible estimates of $r(t)$ from the entire time series.

Analysis of estimated $r(t)$ time series.
To analyze the estimated county- and state-level $r(t)$ time series, we used an autoregressive moving-average (ARMA) time-series model[33]. This statistical modeling approach is parsimonious, robust, and dynamically flexible when fitting linear or approximating nonlinear processes[33]. By
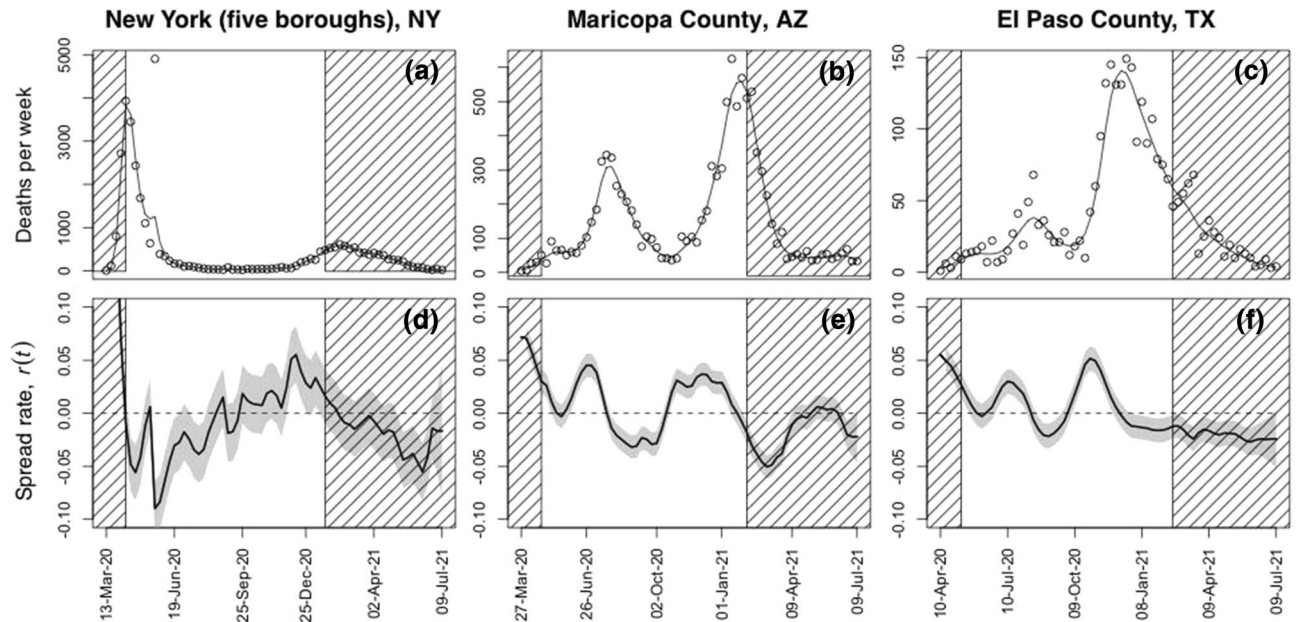
Figure 1. Death data from three illustrative counties in the USA and estimated disease spread rates. Weekly death count data and resulting estimates of $r(t)$ are given for (**a, d**) New York (five boroughs), New York, (**b, e**) Maricopa County, Arizona, and (**c, f**) El Paso County, Texas. In panels (**a–c**), points give the data from Ref.[26], and the lines give the fit from the Kalman smoother. Panels (**d–f**) give the corresponding estimates of the spread rate $r(t)$ from the Kalman smoother, with the shaded band encompassing the 66% confidence interval. Values of $r(t)$ within the hatched region at the start and end of the time series were removed for our analyses to exclude the initial outbreak and possible effects of vaccination on the dynamics.

estimating $r(t)$ separately for each jurisdiction as described in the preceding sub-section, we could allow each jurisdiction to have different statistical attributes, such as how rapidly $r(t)$ changes through time and the magnitude of measurement error. The ARMA time series model then allows us to explore common and contrasting patterns in $r(t)$ among jurisdictions, taking into account spatial autocorrelation that manifests as similar dynamics shown by geographically proximate jurisdictions. We fit a spatial ARMA(2,2) model to both county- and state-level datasets separately, in which each jurisdiction had its own autoregressive coefficients, but all jurisdictions shared the same moving average coefficients, and random errors were assumed to be spatially autocorrelated. We chose the AR order $p = 2$ because it is a parsimonious choice to produce and fit cyclic dynamics[34]. Note, however, that AR(2) dynamics can also be non-cyclic as was the case for several counties and states (cf. Supplementary Fig. S6), and therefore such models allow for more 'dynamical freedom' than using a purely cyclic model. As for the MA order q, we followed the established practice[35] to set $q = p = 2$ to implicitly account for potential measurement error not accounted for while estimating the $r(t)$ time series (see above). We also explored the more-complex model with lags of $q = p = 3$, although this gave results that were indistinguishable from $q = p = 2$, and therefore we only present the results from the simpler model. For more information about our model strategy and model uncertainty, see the Supplementary Information, section *Predictive power, estimation uncertainty, and structural uncertainty* for additional information.

In the spatial ARMA(2,2) model, $r(t)$ in jurisdiction i is given by

$$r_i(t) = b_{0,i} + b_{1,i} r_i(t-1) + b_{2,i} r_i(t-2) + \varepsilon_i(t), \tag{1a}$$

$$\varepsilon_i(t) = a_0 \delta_i(t) + a_1 \delta_i(t-1) + a_2 \delta_i(t-2). \tag{1b}$$

Here, $r_i(t)$ is the spread rate in jurisdiction i for week t, $b_{0,i}$ gives differences in the mean spread rate among jurisdictions, $b_{1,i}$ and $b_{2,i}$ give the jurisdiction-specific AR coefficients for lag-1 and lag-2, a_0, a_1, and a_2 are the MA coefficients for lag-0, lag-1 and lag-2, and $\delta_i(t)$ is a multivariate Gaussian random variable that incorporates spatial correlation. Spatial correlation between two jurisdictions i and j is given by $\text{cor}(\delta_i(t), \delta_j(t)) = (1 - \eta)\exp\left(-\left(\partial_{i,j}\varrho^{-1}\right)^2\right)$, where $\partial_{i,j}$ is the distance between the two jurisdictions, η is the nugget, and ϱ is the range[36]; parameters η and ϱ were estimated along with the AR and MA coefficients.

Cyclic dynamics. The potential cyclicity of the dynamics given by Eq. 1 depends on the estimated ARMA model jurisdiction-specific parameters b_1 and b_2 (e.g. Ref.[34]), where we have dropped the jurisdiction subscript i for clarity. For a stationary oscillatory process, the average cycle length (henceforth period) is $2\pi w^{-1}$, where $\tan(w) = \left(|b_1^2 + 4b_2|\right)^{1/2} b_1^{-1}$. We further use the damping factor d to characterize cyclicity; d scales with the rate at which the amplitude of the cycle decreases over time in the absence of stochasticity. This can be seen in the explicit solution $r(t) = d^{t-1}(r_1\sin(tw) - dr_0\sin((t-1)w))\sin^{-1}(w)$, where r_0 and r_1 are the initial values of $r(t)$

at time point 0 and 1, respectively, and d is the damping factor; for a stationary process, $d < 1$, and values close to zero imply rapid decreases in amplitude. The damping factor can be expressed in terms of the autoregressive lag-2 coefficient as $d^2 = -b_2$.

Predictive power.   To assess predictability, we use the measure predictive power[18], $PP(t)$, which is rooted in information theory. One advantage of working with $PP(t)$ is the ease with which the general framework can be used with linear stochastic systems, like models from the ARIMA family. More fundamentally, predictive power quantifies the amount of information available in a time series for making forecasts, measuring the uncertainty of a prediction. Thus, the focus is not on assessing the ability of specific models to fit the time series and make forecasts. Rather, a predictability measure like $PP(t)$ directly addresses the inherent limit to prediction, in principle valid for all forecasting models. $PP(t)$ is based on the time-dependent variance of the transition distribution (i.e. forecast variance) scaled by the variance of the stationary distribution (i.e. long-term variation) of the ARMA(2,2) process (Supplementary Fig. S3). If both variances are equal, then no information is available for a forecast to be 'better' than a randomly drawn process state according to the stationary distribution, and therefore predictability is said to be lost[12]. Because the transition and stationary distributions are properties of the underlying processes that generate stochastic dynamics, $PP(t)$ gives the theoretical limit of the predictive ability of any model fit to the data.

For a general multivariate Gaussian process, $PP(t)$ is defined as

$$PP(t) := 1 - \det\big(\mathbf{V}(t)\mathbf{V}_\infty^{-1}\big)^{1/2m},\tag{2}$$

where $\det(\cdot)$ is the determinant, $\mathbf{V}(t)$ and $\mathbf{V}_\infty$ are the covariance matrices of the transition and stationary distributions, and m is the dimension; calculation of $\mathbf{V}(t)$ and $\mathbf{V}_\infty$ is outlined in the Supplementary Information, section *Predictive power* for an ARMA(2,2) *process*. Because our ARMA(2,2) model (Eq. 1) is a univariate process, $\mathbf{V}(t)$ and $\mathbf{V}_\infty$ are scalars and $m = 1$. Here, $PP(t)$ can be related to the theoretical limit of forecast accuracy[37]: if $R^2(\tau)$ denotes the coefficient of determination of a predicted value of $r(\tau)$ (τ weeks into the future), then the maximum possible value of $R^2(\tau)$ is $1 - \big(\mathbf{V}(\tau)\mathbf{V}_\infty^{-1}\big) = 1 - (1 - PP(\tau))^2$.

The time dependency of $PP(t)$ implies a decrease in predictability with time, eventually approaching zero (Supplementary Fig. S3). Although the approach to zero is usually defined as the predictability barrier[12], from an empirical perspective, we set the threshold using the link between prediction R^2 and $PP(t)$ as follows. As a rule of thumb[38], values of prediction $R^2 < 0.25$ can be considered as reflecting a very weak match between true and forecasted dynamics. Thus, we set the threshold to compute a predictability barrier as $PP_{\text{lim}} = 1 - (1 - 0.25)^{1/2} = 0.134$. Henceforth, we define predictability barrier as the number of weeks for which $PP(t) = PP_{\text{lim}}$ and where the dynamics beyond this barrier can be considered unpredictable. It is clear that lower values for a limiting prediction R^2 will result in different (i.e. higher) predictability barriers (see the "Results"). Thus, as an additional idea (not pursued further in this study) the often-used root mean square error (RMSE) could be used, which is the standard deviation of the prediction errors[39]. After defining a sensible case-dependent limiting RMSE value, the square of this value could then be used in the expression for prediction R^2 instead of the variance of the transition distribution (see above), which then will allow setting the value of PP_{lim}. As a further note, the computation of $PP(t)$ can also include parameter estimation uncertainty[18]. Nonetheless, because we estimated the ARMA(2,2) parameters from the full time series (40 weeks) and we are dealing with a low-dimensional model, parameter uncertainty is expected to have a marginal effect[40]. Our estimates of $PP(t)$ should be considered optimistic; see the Supplementary Information, section *Predictive power, estimation uncertainty, and structural uncertainty* for additional information.

Finally, to test whether the severity of the initial outbreak (March–early May 2020) affected the ensuing cyclicity and predictability, we used previously estimated values of the basic reproduction number, R_0, from death data at the county and state levels[27,41]; the time period for which these R_0 values were estimated did not overlap with the time series used in the present study. The method for estimating these R_0 values used the observed death counts and a statistical state-space modeling approach similar to our computation of $r(t)$ in the present study. Also, the estimation of the R_0 values in the previous studies was designed to factor out the effects of the timing of epidemic onset (higher spread rates occurred earlier in the epidemic) and population size (to correct for bias in the estimates of R_0). Nonetheless, the estimates of the R_0 values are directly comparable to $r(t)$; they use the same type of data and methodology, but characterize different periods of the pandemic and different dynamical characteristics; see Ref.[27] for further technical details.

Simulations.   To help interpret the $r(t)$ time series and investigate possible mechanisms underlying their cyclicity, we used a stochastic, discrete-time, age-of-infection Susceptible-Infectious-Removed (SIR) model, parameterized with published results[27]. This simulation model tracks the epidemic on a daily time scale and explicitly includes the time period from infection to subsequent transmission (infectiousness), and from infection to death when the disease is reported. We modified the published model to explicitly separate a constitutive disease reproduction number, henceforth R_{const}, from dynamic changes in the transmission rate that depend on the death count two weeks previously; therefore, R_{const} has a fixed transmission rate (Eq. 3). In this way, we mimicked a susceptible population becoming aware of increases in the death toll and, following a 2-week delay for reporting and media attention, taking protective measures (individual behavioral responses and/or NPIs)[23,24,42]. We set the transmission rate to

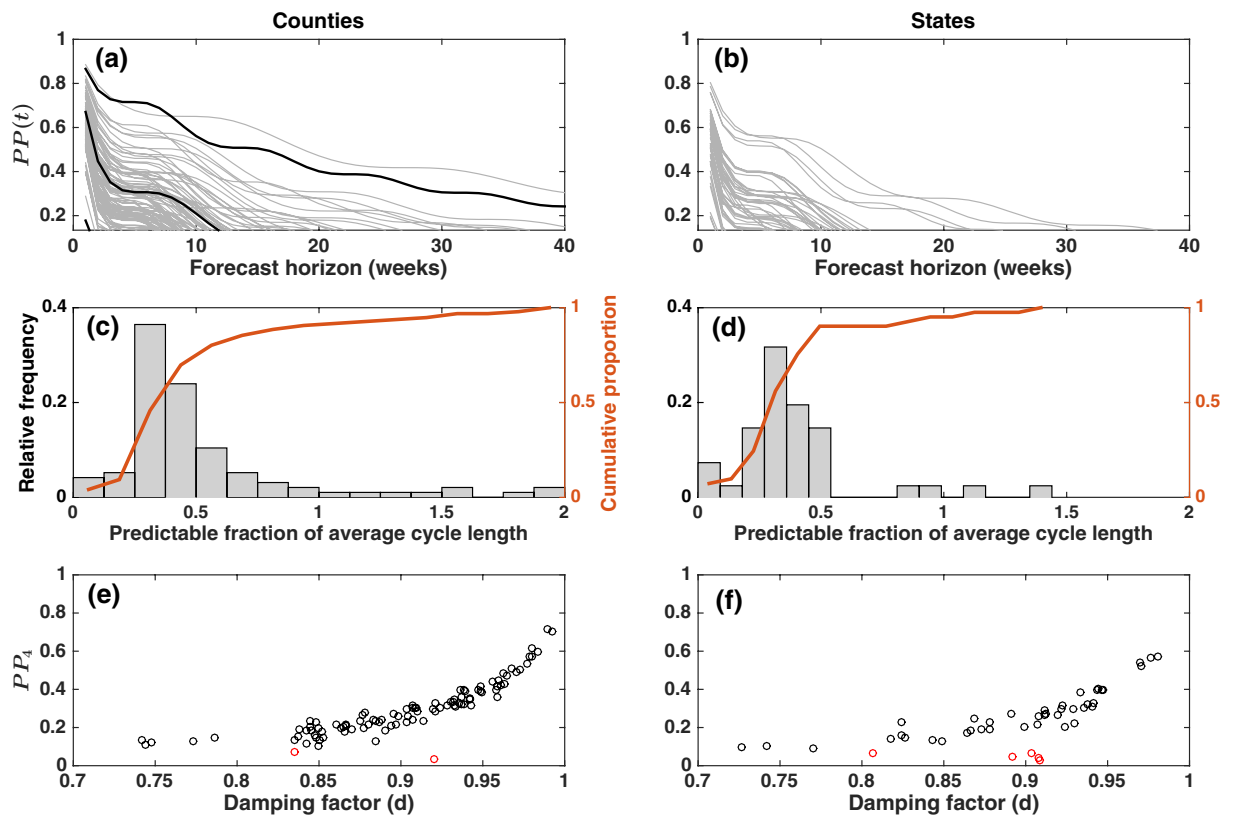$$\beta(t) = \beta_{\text{const}}(1 + \omega D(t - 2))^{-1},\tag{3}$$

Figure 2. Predictive power and cyclicity. Predictive power, $PP(t)$ (Eq. 2), for (**a**) 98 counties and (**b**) 46 conterminous states with stationary dynamics for a forecast horizon ranging from 1 to 40 weeks. The lowest value of the y-axes in (**a**, **b**) is the threshold $PP_{lim} = 0.134$, used to compute the predictability barrier (cf. Supplementary Fig. S4). The three highlighted counties in (**a**) are those in Figs. 1 and 3: New York (five boroughs), New York; Maricopa County, Arizona; El Paso County, Texas. Panels (**c**, **d**) show the distribution of the estimated predictability barriers of jurisdictions with cyclic dynamics (cf. panels (**a**, **b**) and Supplementary Fig. S4) as a fraction of the average cycle length (counties: 23 weeks, states: 26 weeks), along with the cumulative proportion of jurisdictions. Finally, for both (**e**) counties and (**f**) states, PP_4 was closely associated with the strength of cyclicity as measured by the damping factor d; the latter corresponds to the characteristic return time[51] for non-cyclic jurisdictions (red dots).

where β_{const} is the transmission rate corresponding to R_{const}, $D(t-2)$ is the number of deaths two weeks previously, and ω scales how rapidly the transmission rate decreases with increases in $D(t-2)$. We selected this functional form to mimic the cyclicity in the observed data, although similar disease dynamics may be generated using other functions that decrease with $D(t-2)$. Our modeling approach is similar to that used by Weitz et al.[25], although our model explicitly incorporates the dependence of transmission and death on the number of days since infection, making it possible to compare our simulation results with real data. For further simulation details, see the Supplementary Information, section *Simulation model*.

The simulation model is built on the hypothesis that cyclicity is determined by differences in the constitutive and/or dynamic components of the transmission rate among jurisdictions. Our analyses, however, do not test this hypothesis directly. Instead, by comparing the simulated and real dynamics, we ask whether the hypothesis is plausible.

Results

Predictability and cyclicity at the county and state levels.

Predictability measured by $PP(t)$ varied substantially among counties and states (Fig. 2). For example, at the county level and for four-week-ahead forecasts, PP_4 ranged from 0.03 to 0.72. This among-jurisdiction variation in $PP(t)$ for any week t reflected high variation in the predictability barrier (Fig. 2a,b, Supplementary Fig. S4): counties had a median of 9 weeks (interquartile range 7–12 weeks), and states had a median of 8 weeks (5–11). PP_4—chosen to reflect the empirically found barrier of four weeks (see the "Introduction")—characterizes the variation in predictability barrier among jurisdictions (Supplementary Fig. S5), and therefore we focus on PP_4 throughout most of the remaining analyses.

Of the 100 counties and 49 states, 96 and 41 showed cyclic dynamics in the stationary domain (Supplementary Fig. S6). The estimated period was similar at the county and state levels (Supplementary Fig. S7a,b): counties had a median of 23 weeks (interquartile range 20–29), and states had a median of 26 weeks (20–33). The damping factor (d) was also similar (Supplementary Fig. S7c,d): counties had median $d = 0.91$ (0.85–0.96), and states had median $d = 0.91$ (0.83–0.94).
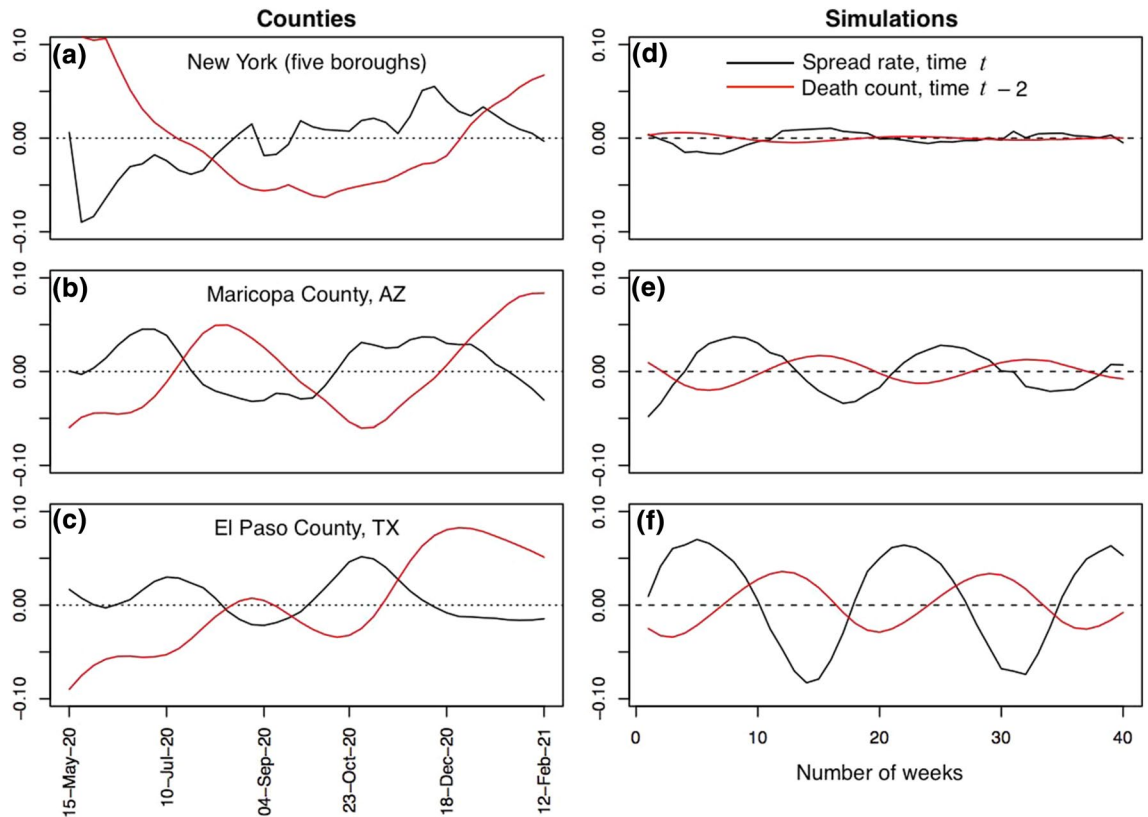
Figure 3. Comparison between $r(t)$ estimated for three illustrative counties and for three simulated populations. The three counties are those in Fig. 1: (**a**) New York (five boroughs), New York, for which $PP_4 = 0.03$, (**b**) Maricopa County, Arizona, with $PP_4 = 0.32$, and (**c**) El Paso County, Texas, with $PP_4 = 0.72$. For the simulations, β_{const} was selected to give three values of the constitutive reproduction number R_{const} (Eq. 3): (**d**) $R_{const} = 1.4$, $PP_4 = 0.11$, (**e**) $R_{const} = 1.6$, $PP_4 = 0.28$, and (**f**) $R_{const} = 1.8$, $PP_4 = 0.55$. In all panels, black lines give the estimates of $r(t)$ and red lines give the z-transformed log number of deaths per week incorporating a 2-week time lag.

Expressing the predictability barrier as a fraction of the median period (23 weeks and 26 weeks, see above) shows that for half the counties with stationary cyclic dynamics, at most 40% of a cycle is predictable, while at the state level it is 35% (Fig. 2c,d). Furthermore, only 10% of counties and 5% of states had a fully predictable cycle ('wave') or beyond. Results of predictability barriers presented so far are based on a predictability threshold (PP_{lim}) computed using a limiting prediction R^2 value of 0.25 (see "Materials and methods"). In the Supplementary Table S1 we compare these results with results based on a (much) lower limiting prediction R^2 of 0.10. As expected, predictability barrier values increase, but not dramatically so: for example, half of all counties and states still have only approximately 50% of the respective median period predictable. Nonetheless, given the dependence of the predictability barrier on a preset threshold, as justified above (cf. Supplementary Fig. S5) we focus on PP_4 throughout the remaining analyses.

Exploring cyclicity further, we found a strong association between predictability and damping factor (Fig. 2e,f) (counties: Spearman's $\varrho = 0.83$, $P < 10^{-10}$; states: $\varrho = 0.52$, $P = 0.0001$). This result is not a mathematical inevitability: for example, near-random-walk dynamics are non-cyclic yet still imply high damping factors. In contrast to this association, we could not find a significant relationship between predictability and period (Supplementary Fig. S8), and therefore we will use the damping factor as a measure of cyclicity to investigate what causes the joint variation in cyclicity and predictability.

Simulation results. The simulation model mimics the cyclic dynamics shown in the data (Fig. 3). Increases in cyclicity and predictability in the simulations are generated by increasing the constitutive reproduction number, R_{const}. Because higher R_{const} values correspond to higher maximum values of $r(t)$, more pronounced cyclicity and increased predictability occur when there is greater potential for rapid increases in disease spread rates. In the specific model realizations, increasing the R_{const} value from 1.4 (Fig. 3d) to 1.8 (Fig. 3f) increases PP_4 from 0.11 to 0.55.

To compare with the county-level data, we simulated time series of 40 weeks using values of R_{const} randomly distributed between 1.4 and 1.8 (Fig. 4). Analyzing the simulated data in the same way as the real data, these simulations spanned the range of PP_4 observed in the county data (Fig. 4a). In the simulations, the association between the damping factor d and PP_4 (Fig. 4b) was very close to that found for the county data (Fig. 4e). The
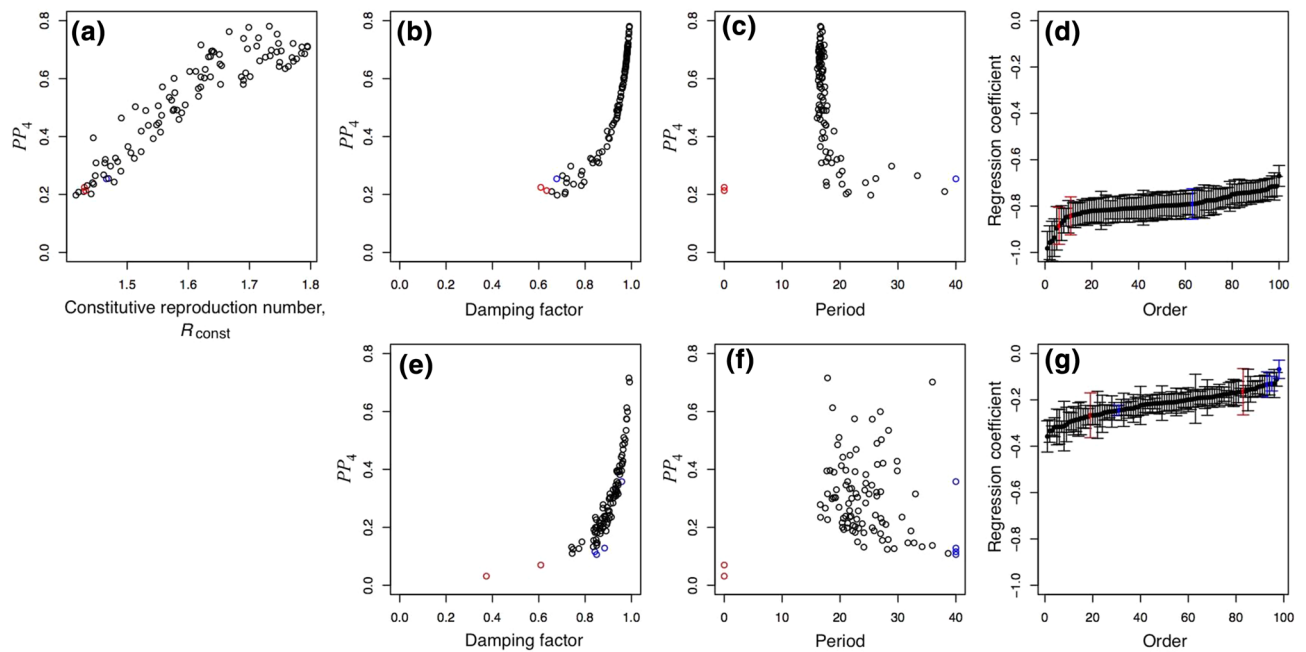
Figure 4. Comparison between dynamical characteristics of $r(t)$ for simulated and real time series. One-hundred time series of $r(t)$ were simulated for randomly selected values of the transmission rate β_{const} (Eq. 3) to give values of the constitutive reproduction number (R_{const}) ranging from 1.4 to 1.8. (**a**) Predictability measured by PP_4 increased with R_{const} implying that simulations in which the potential for increases in the spread rate were greatest (those with highest R_{const}) had the most predictable dynamics. Simulated time series that were identified as non-cyclic are shown in red, and time series with period > 40 weeks are shown in blue, respectively. The association between predictability and cyclicity for (**b**) the simulated data was similar to that for (**e**) the real data. For simulated counties, (**c**) the period showed less variation than (**f**) the real data, although the overall relationship with PP_4 was similar. No simulated time series was non-stationary, and the two non-stationary county time series are excluded because PP is undefined. Finally, (**d**) and (**g**) show the ranked regression coefficients for the effect of $D(t-2)$ on $r(t)$ for the simulated counties and county data, where bars give standard errors.

periods estimated from the simulated data were less variable than for the real data, although most fell between 20 and 30 weeks (Fig. 4c,f).

The key feature of the simulations generating cycles is the decrease in the transmission rate caused by increases in the death count two weeks beforehand (Eq. 3). This feature of the simulation can be recovered statistically from the simulated time series by performing a conditional least-squares regression of $r(t)$ against $r(t-1)$ and $D(t-2)$. For the 100 simulated counties, the regression coefficients ranged between -1 and -0.4 (Fig. 4d). For the county data, these regression coefficients ranged between -0.4 and -0.05 (Fig. 4g), and all but one (for a non-cyclic time series) are statistically significantly below zero ($P < 0.05$).

R_0 and variation in predictability.
At both the county and state levels, the R_0 values and PP_4 were strongly negatively associated (Fig. 5a,b; counties: Spearman's $\varrho = -0.63$, $P < 10^{-10}$; states: $\varrho = -0.52$, $P = 0.001$): more severe initial outbreaks were followed by disease spread dynamics with lower predictability. This is the opposite pattern from what would be expected if high R_0 values were followed by high constitutive reproduction number values (R_{const}); in the simulations, higher R_{const} values were associated with higher PP_4 (Fig. 4a). These results imply that higher R_0 values gave rise to ensuing dynamics with lower R_{const} values, suggesting that populations were constitutively more cautious in counties and states that had experienced a severe COVID-19 outbreak at the start of the pandemic.

Figure 5c overlays county estimates of PP_4 on a map of the county estimates of R_0 values from the initial outbreaks. A cluster of counties with low PP_4 occurs along the northeastern coast where R_0 values were high, while counties with high PP_4 and pronounced cyclicity occur in southern states and in California.

Discussion
The COVID-19 pandemic has stimulated the development of numerous quantitative models to help understand and forecast disease dynamics, and to assist public health decision-making (e.g. Ref.[11,19,43]). Rather than develop methods for making predictions, in this study we have focused on the inherent unpredictability of COVID-19 dynamics. Our goals have been both to address the limits to which predictions are possible for communicable diseases like COVID-19, and to understand the dynamical characteristics of epidemics that make predictions more or less accurate.
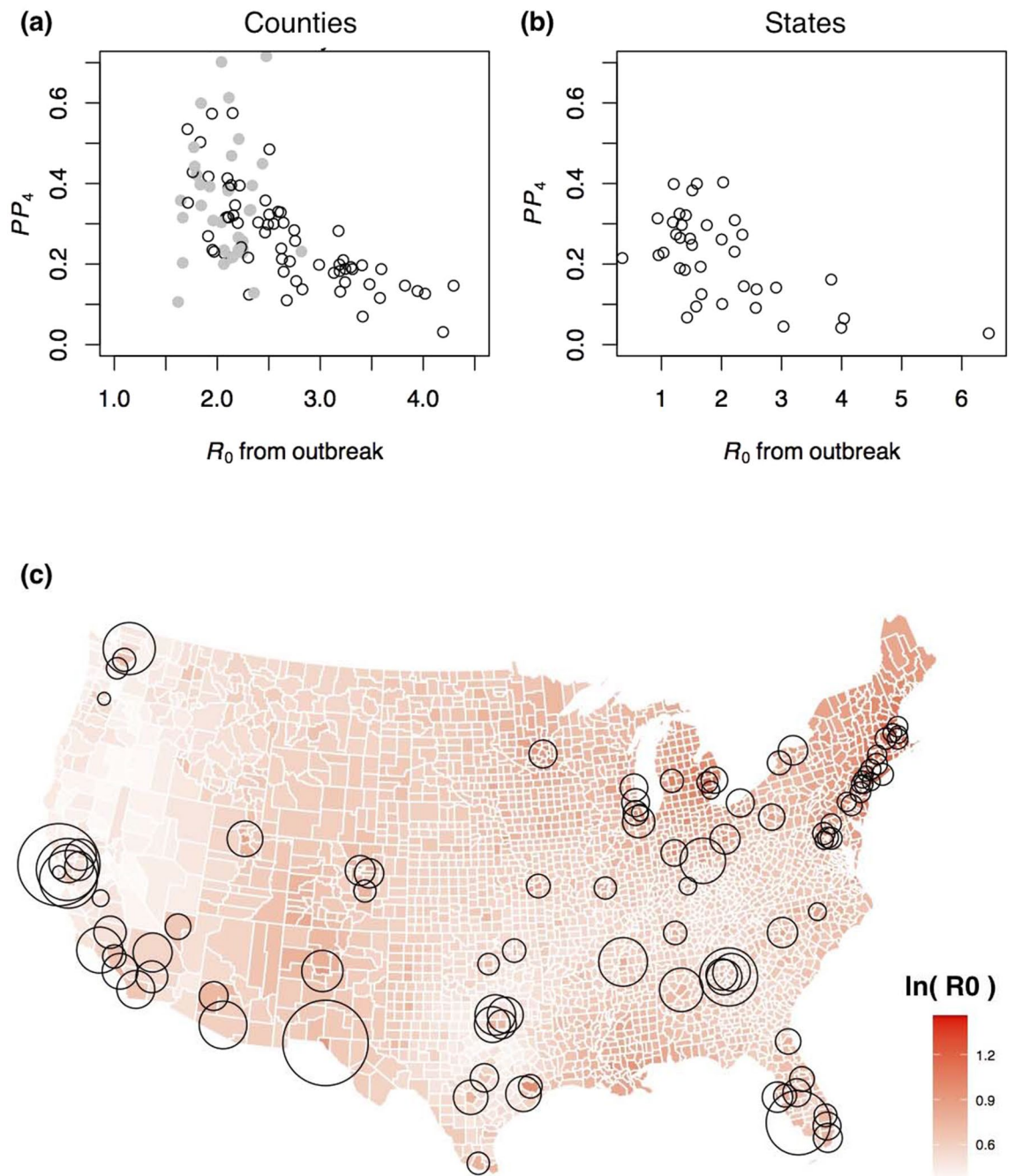
Figure 5. Effect of R_0 on predictability. (**a**, **b**) Predictability as measured by PP_4 is plotted against the R_0 values estimated at the onset of the pandemic from death count records[27,41]. For county-level data, R_0 values were computed for 124 counties, and those counties overlapping with the counties in the present study are shown as open points; the gray points correspond to counties for which values of R_0 were interpolated using population density and geographical location. (**c**) The map gives the log-transformed county-level estimates of R_0, originally ranging from 1.3 to 4.5 (white to dark red; Ref.[27]), with county-level estimates of PP_4 depicted by circles, with circle diameter from smallest to largest corresponding to PP_4 from 0.03 to 0.72.

We found considerable variation in predictability among jurisdictions (Fig. 2, Supplementary Fig. S4), as also found by Scarpino and Petri[14]. In contrast to these authors[14], however, we found that for the majority of analyzed counties and states, the predictable fraction of a cycle (that is, an outbreak in Ref.[14]) is much less than one (Fig. 2). Our estimated cycle lengths are in good agreement with previous findings[9,10]. In addition, we show that predictability is strongly related to the rate at which cycles are damped, with weakly damped cycles giving regular patterns in the data that allow predictions: this rate of cycle damping has been largely neglected in previous analyses. Finally, we show that protective measures against COVID-19 can reduce both the cyclicity

and predictability of disease dynamics. Thus, variation in cyclicity and predictability among jurisdictions gives valuable information about factors governing the dynamics of COVID-19.

In analyses of forecast accuracy, single studies and reviews of the many studies forecasting COVID-19 dynamics have focused on identifying the best forecasting methods (e.g. Ref.[11,17]). Our analyses of inherent unpredictability focus on how much information is available in a time series, rather than the ability of a model to fit the time series and make forecasts. Therefore, our estimates of the limits to forecasts in principle should apply to all forecasting models. Furthermore, our demonstration of the high variation in predictability among time series from different counties and states in the USA implies that the ability to forecast COVID-19 likely depends more on the dynamics in a particular dataset than on the forecasting methods used.

Our simulation model showed that cyclic dynamics similar to those observed in the county and state data can be mimicked when changes in the transmission rate occur as a 2-week delayed response to increases in the death toll. We acknowledge that this is not categorical evidence that time-delayed changes in the transmission rate in response to death counts are responsible for the cycles, because any form of cyclicity in $D(t)$ will drive cyclicity in $r(t)$. Nonetheless, this pattern is consistent with the hypothesis under which the simulation model was built. The simulation model shows the plausibility of the hypothesis that more pronounced cyclicity occurs in jurisdictions with higher constitutive reproduction number values (R_{const}), because a higher R_{const} allows more rapid changes in the transmission rate that are necessary to generate cycles. Finally, jurisdictions that experienced severe outbreaks at the onset of the pandemic, measured by high values of R_0 before widespread public protective measures were put in place, had less cyclic and less predictable COVID-19 dynamics in the subsequent period before vaccination became widespread. The association between a high R_0 value and lack of predictability suggests that a severe initial outbreak led to high levels of constitutive protective measures which individuals took to reduce disease transmission. Moreover, the variation in predictability had a clear geographical pattern, with many counties having unpredictable dynamics occurring in the Northeast (Fig. 5).

The hypothesis embodied by our simulation model is that cyclicity arises from protective measures people take in response to rising death tolls (cf. Ref.[24]), that is, a negative feedback loop much like "predator–prey" dynamics in ecology which has recently attracted increased attention in epidemiology (Ref.[23], and references therein). Because death tolls are highly correlated with case counts, human responses could equally depend on the awareness of rising cases, reports in the media, word-of-mouth, etc. Maps of current cases and deaths from COVID-19 were publicly available throughout the time period we analyzed, and reports of case counts occurred regularly in the news. Some responses to increased spread of COVID-19 were taken by policy-makers, such as mask mandates and restaurant closures. Other responses were taken by individuals to reduce contact and abide mandates. We have shown that if the 'background' constitutive transmission rate of COVID-19 is high, then the human response to increasing disease spread will generate pronounced cyclic dynamics. In contrast, if the constitutive transmission rate is kept low, then cycles do not appear, because the disease dynamics are not as responsive to changes in protective measures. This implies that lack of cyclicity and predictability are caused when people continuously take greater precautions against COVID-19, rather than show an on-and-off response to changes in death tolls or case counts.

There has been considerable research effort to assess attitudes, such as surveys on mask use[44] and vaccination hesitancy[45], and to identify effective proxies of protective behaviors, such as analyses of government policies[28] and changes in individual movement patterns using cell-phone signals[46]. While acknowledging the value of these studies, our approach of analyzing the dynamics of COVID-19 focuses on the effects of protective behaviors, rather than the protective behaviors themselves. Even though our approach cannot make a mechanistic link between behaviors and dynamics, it nonetheless gives insight into differences in how COVID-19 was experienced by different jurisdictions.

Our explanation for the joint variation in cyclicity and predictability is a hypothesis that is consistent with our statistical evidence. Direct evidence is a challenge, however, because variation among jurisdictions in the constitutive protective measures that individuals take are hard to document. Nonetheless, the remarkable negative association between predictability and R_0 (Fig. 5) suggests differences in personal protective measures among jurisdictions. Before performing our analyses, we hypothesized that R_0 values would be positively associated with predictability, because a high R_0 value implies the potential for rapid increases in disease spread if protective measures were dropped. Our finding of a negative association suggests that populations experiencing severe initial outbreaks saw a fundamental shift in later transmission rates. An alternative explanation for this shift is that the initial outbreak generated sufficient acquired immunity to reduce future transmission rates[10]. Arguing against this explanation, however, is that during the period we analyzed the number of COVID-19 cases as a proportion of the population ranged from 1 to 14% among counties and 2–13% among states. Furthermore, there was no relationship between the cumulative per capita number of cases and PP_4 for either county (Spearman's $\varrho = 0.12$, $P = 0.22$) or state data ($\varrho = 0.23$, $P = 0.11$). Even though cases were likely under-reported, serological studies show that, for example, the proportion of the adult population in New York City having contracted COVID-19 between 19 April and 5 July, 2020, was approximately 20%—similar results have been found for metropolitan France (approximately 15% of adults by January 2021)—which is likely not high enough to affect the subsequent predictability of the dynamics[47,48]. It is also possible that cyclicity was driven by successive SARS-CoV-2 variants each with higher transmission rates[22]. While different variants are associated with differences in R_0 among jurisdictions at the start of the pandemic[27], and successive variants were more transmissible[49], the successive variants spread geographically quickly throughout the conterminous USA. While new variants might have added to the broad pattern of cyclicity of COVID-19 in the USA, we cannot think of how new variants could explain the negative association between R_0 values and subsequent cyclicity. Given that acquired immunity and SARS-CoV-2 variants are unlikely explanations for the negative association between R_0 and predictability across jurisdictions, our best alternative is changes in protective measures taken by individuals.

What are the implications of our findings for decision-making in public health emergencies? The USA experienced repeated waves of COVID-19 after the initial spread of the pandemic, and these waves caused large numbers of infections and deaths. Nonetheless, after the initial rapid outbreaks, the spread rates were lower (compare the results in Ref.[27] to Supplementary Fig. S1). This suggests that steps taken by policy-makers and individuals to reduce transmission rates—such as mask wearing, social distancing, and other NPIs—were effective. Indeed, the lack of predictability can be viewed as a consequence of the successful maintenance of low transmission rates. If COVID-19 spread rates are predictable, it means that protective measures have been dropped and therefore have to be restarted. Although the consequence of a population taking continuous protective measures is lack of predictability, lack of predictability itself is an indicator of effective transmission management. Our results further indicate that one of the first epidemic-related metrics computed at the early stages of an epidemic, namely R_0, allows anticipating the predictability of the ensuing dynamics (Fig. 5). For outbreaks of newly emerged diseases this information could be complemented by jurisdiction-specific data indicating how well NPIs in the past have been successful, in terms of swift implementation and adherence by the population (e.g. Ref.[28,43]): this would give information about how strongly protective measures will affect disease dynamics and consequently their predictability. Finally, all our results are similar at the county and state levels, implying that at the onset of outbreaks, information from different jurisdictional levels could be helpful to gauge the limit to forecasting accuracy.

The human response to disease spread likely affects its predictability, and a pandemic might be similar to stock markets in which unpredictability is generated by human behavior[50]. We should anticipate that future pandemics will be similarly unpredictable if they elicit widespread behaviors to reduce transmission. Unpredictability is just a by-product of positive steps that people take to protect themselves and others.

Data availability

References

1. Morens, D. M. & Fauci, A. S. Emerging infectious diseases: Threats to human health and global stability. *PLoS Pathog.* **9**, e1003467 (2013).
2. Bloom, D. E. & Cadarette, D. Infectious disease threats in the twenty-first century: Strengthening the global response. *Front. Immunol.* **10**, 549 (2019).
3. Mora, C. *et al.* Over half of known human pathogenic diseases can be aggravated by climate change. *Nat. Clim. Change* https://doi.org/10.1038/s41558-022-01426-1 (2022).
4. Molnár, O., Hoberg, E., Trivellone, V., Földvári, G. & Brooks, D. R. *The 3P Framework—A Comprehensive Approach to Coping with the Emerging Infectious Disease Crisis.* https://doi.org/10.22541/au.166176189.90109497/v1 (2022).
5. Lipsitch, M., Finelli, L., Heffernan, R. T., Leung, G. M. & Redd, S. C. Improving the evidence base for decision making during a pandemic: The example of 2009 influenza A/H1N1. *Biosecur. Bioterror.* **9**, 89–115 (2011).
6. Lutz, C. S. *et al.* Applying infectious disease forecasting to public health: A path forward using influenza forecasting examples. *BMC Public Health* **19**, 1659 (2019).
7. Doms, C., Kramer, S. C. & Shaman, J. Assessing the use of influenza forecasts and epidemiological modeling in public health decision making in the United States. *Sci. Rep.* **8**, 12406 (2018).
8. Viboud, C. & Vespignani, A. The future of influenza forecasts. *PNAS* **116**, 2802–2804 (2019).
9. Huang, J. *et al.* The oscillation-outbreaks characteristic of the COVID-19 pandemic. *Natl. Sci. Rev.* **8**, nwab100 (2021).
10. Mitchell, R. N. & Zhang, J. Four-month intrinsic viral cycle in COVID-19. *Innovation* **3**, 100196 (2022).
11. Gnanvi, J. E., Salako, K. V., Kotanmi, G. B. & Glèlè Kakaï, R. On the reliability of predictions on Covid-19 dynamics: A systematic and critical review of modelling techniques. *Infect. Dis. Modell.* **6**, 258–272 (2021).
12. DelSole, T. Predictability and information theory. Part I: Measures of predictability. *J. Atmos. Sci.* **61**, 2425–2440 (2004).
13. Fernandes, L. H. S., Araujo, F. H. A., Silva, M. A. R. & Acioli-Santos, B. Predictability of COVID-19 worldwide lethality using permutation-information theory quantifiers. *Results Phys.* **26**, 104306 (2021).
14. Scarpino, S. V. & Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **10**, 898 (2019).
15. Reich, N. G. *et al.* Accuracy of real-time multi-model ensemble forecasts for seasonal influenza in the U.S. *PLOS Comput. Biol.* **15**, e1007486 (2019).
16. Gordeev, D., Singer, P., Michailidis, M., Müller, M. & Ambati, S. Backtesting the predictability of COVID-19. arXiv:2007.11411 [physics, q-bio] (2020).
17. Cramer, E. Y. *et al.* Evaluation of individual and ensemble probabilistic forecasts of COVID-19 mortality in the US. *medRxiv* https://doi.org/10.1101/2021.02.03.21250974 (2021).
18. Schneider, T. & Griffies, S. M. A conceptual framework for predictability studies. *J. Clim.* **12**, 3133–3155 (1999).
19. Duarte, P. & Riveros-Perez, E. Understanding the cycles of COVID-19 incidence: Principal component analysis and interaction of biological and socio-economic factors. *Ann. Med. Surg.* **66**, 102437 (2021).
20. Keeling, M. J. & Rohani, P. *Modeling Infectious Diseases in Humans and Animals* (Princeton University Press, 2008).
21. Bozzuto, C. & Canessa, S. Impact of seasonal cycles on host-pathogen dynamics and disease mitigation for *Batrachochytrium salamandrivorans*. *Glob. Ecol. Conserv.* **17**, e00551 (2019).
22. Callaway, E. Are COVID surges becoming more predictable? New Omicron variants offer a hint. *Nature* **605**, 204–206 (2022).
23. Just, W., Saldaña, J. & Xin, Y. Oscillations in epidemic models with spread of awareness. *J. Math. Biol.* **76**, 1027–1057 (2018).
24. Glaubitz, A. & Fu, F. Oscillatory dynamics in the dilemma of social distancing. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **476**, 20200686 (2020).
25. Weitz, J. S., Park, S. W., Eksin, C. & Dushoff, J. Awareness-driven behavior changes can shift the shape of epidemics away from peaks and toward plateaus, shoulders, and oscillations. *Proc. Natl. Acad. Sci.* **117**, 32764–32771 (2020).
26. CSSEGISandData. *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University* (2021).
27. Ives, A. R. & Bozzuto, C. Estimating and explaining the spread of COVID-19 at the county level in the USA. *Commun. Biol.* **4**, 1–9 (2021).

28. Dey, T. *et al.* Lag time between state-level policy interventions and change points in COVID-19 outcomes in the United States. *PATTER* **2**, 100306 (2021).
29. Anderson, R. M., Grenfell, B. T. & May, R. M. Oscillatory fluctuations in the incidence of infectious disease and the impact of vaccination: Time series analysis. *Epidemiol. Infect.* **93**, 587–608 (1984).
30. Carletti, T., Fanelli, D. & Piazza, F. COVID-19: The unreasonable effectiveness of simple models. *Chaos Solitons Fract.: X* **5**, 100034 (2020).
31. Park, S. W. *et al.* Reconciling early-outbreak estimates of the basic reproductive number and its uncertainty: Framework and applications to the novel coronavirus (SARS-CoV-2) outbreak. *J. R. Soc. Interface* **17**, 20200144 (2020).
32. Durbin, J. & Koopman, S. J. *Time Series Analysis by State Space Methods* (Oxford University Press, 2012).
33. Box, G. E. P., Jenkins, G. M., Reinsel, G. C. & Ljung, G. M. *Time Series Analysis: Forecasting and Control* (Wiley, 2015).
34. Royama, T. *Analytical Population Dynamics* (Springer, 1992). https://doi.org/10.1007/978-94-011-2916-9.
35. Staudenmayer, J. & Buonaccorsi, J. P. Measurement error in linear autoregressive models. *J. Am. Stat. Assoc.* **100**, 841–852 (2005).
36. Cressie, N. *Statistics for Spatial Data* (Wiley, 2015).
37. Ives, A. R. R^{2}s for correlated data: Phylogenetic models, LMMs, and GLMMs. *Syst. Biol.* **68**, 234–251 (2019).
38. Hair, J. F. Jr., Sarstedt, M., Hopkins, L. & Kuppelwieser, V. G. Partial least squares structural equation modeling (PLS-SEM): An emerging tool in business research. *Eur. Bus. Rev.* **26**, 106–121 (2014).
39. Hyndman, R. J. & Athanasopoulos, G. *Forecasting: Principles and Practice* (OTexts, 2021).
40. Lütkepohl, H. *New Introduction to Multiple Time Series Analysis* (Springer, 2005).
41. Ives, A. R. & Bozzuto, C. State-by-State estimates of R0 at the start of COVID-19 outbreaks in the USA. *MedRXiv* https://doi.org/10.1101/2020.05.17.20104653 (2020).
42. Dönges, P. *et al.* Interplay between risk perception, behavior, and COVID-19 spread. *Front. Phys.* **10**, 842180 (2022).
43. Flaxman, S. *et al. Report 13: Estimating the number of infections and the impact of non-pharmaceutical interventions on COVID-19 in 11 European countries. 35.* http://spiral.imperial.ac.uk/handle/10044/1/77731. https://doi.org/10.25561/77731 (2020).
44. Rader, B. *et al.* Mask-wearing and control of SARS-CoV-2 transmission in the USA: A cross-sectional study. *Lancet Digit. Health* **3**, e148–e157 (2021).
45. Khubchandani, J. *et al.* COVID-19 vaccination hesitancy in the United States: A rapid national assessment. *J. Community Health* **46**, 270–277 (2021).
46. Levin, R., Chao, D. L., Wenger, E. A. & Proctor, J. L. Insights into population behavior during the COVID-19 pandemic from cell phone mobility data and manifold learning. *Nat. Comput. Sci.* **1**, 588–597 (2021).
47. Hozé, N. *et al.* Monitoring the proportion of the population infected by SARS-CoV-2 using age-stratified hospitalisation and serological data: A modelling study. *Lancet Public Health* **6**, e408–e415 (2021).
48. Stadlbauer, D. *et al.* Repeated cross-sectional sero-monitoring of SARS-CoV-2 in New York City. *Nature* **590**, 146–150 (2021).
49. Panovska-Griffiths, J. *et al.* Statistical and agent-based modelling of the transmissibility of different SARS-CoV-2 variants in England and impact of different interventions. *Philos. Trans. R. Soc. A: Math. Phys. Eng. Sci.* **380**, 20210315 (2022).
50. Sornette, D. *Why Stock Markets Crash* (2017).
51. Ives, A. R., Abbott, K. C. & Ziebarth, N. L. Analysis of ecological time series with ARMA(p, q) models. *Ecology* **91**, 858–871 (2010).

Acknowledgements

Author contributions

C.B. and A.R.I. designed the study and methodology; C.B. and A.R.I. performed the analyses; C.B. and A.R.I. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-40990-0.

Correspondence and requests for materials should be addressed to C.B.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.