Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period

Stephen M. Kissler*, Christine Tedijanto*, Edward Goldstein†, Yonatan Grad†, Marc Lipsitch††

*Department of Immunology and Infectious Diseases, Harvard T.H. Chan School of Public Health, Boston, MA, USA. †Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA.

*These authors contributed equally to this work and are co-senior authors.
†These authors contributed equally to this work.
‡Corresponding author. Email: mlipsitc@hsph.harvard.edu (M.L.); yggrad@hsph.harvard.edu (Y.H.G.)

It is urgent to understand the future of severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) transmission. We used estimates of seasonality, immunity, and cross-immunity for betacoronaviruses OC43 and HKU1 from time series data from the USA to inform a model of SARS-CoV-2 transmission. We projected that recurrent wintertime outbreaks of SARS-CoV-2 will probably occur after the initial, most severe pandemic wave. Absent other interventions, a key metric for the success of social distancing is whether critical care capacities are exceeded. To avoid this, prolonged or intermittent social distancing may be necessary into 2022. Additional interventions, including expanded critical care capacity and an effective therapeutic, would improve the success of intermittent distancing and hasten the acquisition of herd immunity. Longitudinal serological studies are urgently needed to determine the extent and duration of immunity to SARS-CoV-2. Even in the event of apparent elimination, SARS-CoV-2 surveillance should be maintained since a resurgence in contagion could be possible as late as 2024.

The ongoing severe acute respiratory syndrome–coronavirus 2 (SARS-CoV-2) pandemic has caused nearly 500,000 detected cases of coronavirus disease 2019 (COVID-19) illness and claimed over 20,000 lives worldwide as of 26 Mar 2020 (1). Experience from China, Italy, and the United States demonstrates that COVID-19 can overwhelm even the healthcare capacities of well-resourced nations (2–4). With no pharmaceutical treatments available, interventions have focused on contact tracing, quarantine, and social distancing. The required intensity, duration, and urgency of these responses will depend both on how the initial pandemic wave unfolds and on the subsequent transmission dynamics of SARS-CoV-2. During the initial pandemic wave, many countries have adopted social distancing measures, and some, like China, are gradually lifting them after achieving adequate control of transmission. However, to mitigate the possibility of resurgence of infection, prolonged or intermittent periods of social distancing may be required. After the initial pandemic wave, SARS-CoV-2 might follow its closest genetic relative, SARS-CoV-1, and be eradicated by intensive public health measures after causing a brief but intense epidemic (5). Increasingly, public health authorities consider this scenario unlikely (6). Alternatively, the transmission of SARS-CoV-2 could resemble that of pandemic influenza by circulating seasonally after causing an initial global wave of infection (7). Such a scenario could reflect the previous emergence of known human coronaviruses from zoonotic origins e.g. human coronavirus (HCoV) OC43 (8). Distinguishing between these scenarios is key for formulating an effective, sustained public health response to SARS-CoV-2.

The pandemic and post-pandemic transmission dynamics of SARS-CoV-2 will depend on factors including the degree of seasonal variation in transmission, the duration of immunity, and the degree of cross-immunity between SARS-CoV-2 and other coronaviruses, as well as the intensity and timing of control measures. SARS-CoV-2 belongs to the betacoronavirus genus, which includes the SARS-CoV-1 coronavirus, MERS coronavirus, and two other human coronaviruses, HCoV-OC43 and HCoV-HKU1. The SARS-CoV-1 and MERS coronaviruses cause severe illness with approximate case fatality rates of 9 and 36% respectively, but the transmission of both has remained limited (9). HCoV-OC43 and HCoV-HKU1 infections may be asymptomatic or associated with mild to moderate upper respiratory tract illness; these HCoVs are considered the second most common cause of the common cold (9). HCoV-OC43 and HCoV-HKU1 cause annual wintertime outbreaks of respiratory illness in temperate regions (10, 11), suggesting that wintertime climate and host behaviors may facilitate transmission, as is true for influenza (12–14). Immunity to HCoV-OC43 and HCoV-HKU1 appears to wane appreciably within one year (15), while SARS-CoV-1 infection can induce longer-lasting immunity (16). The betacoronaviruses can induce immune responses against one another: SARS-CoV-1 infection can generate neutralizing antibodies
against HCoV-OC43 (16) and HCoV-OC43 infection can generate cross-reactive antibodies against SARS-CoV-1 (17). While investigations into the spectrum of illness caused by SARS-CoV-2 are ongoing, recent evidence indicates the majority of cases experience mild to moderate illness with more limited occurrence of severe lower respiratory infection (18). Current COVID-19 case fatality rates are estimated to lie between 0.6% and 3.5% (19, 20), suggesting lower severity than SARS-CoV-1 and MERS but higher severity than HCoV-OC43 and HCoV-HKU1. The high infectiousness near the start of the SARS-CoV-2 transmission is considerably harder to control with case-based interventions such as intensive testing, isolation and tracing, compared to SARS-CoV-1 and MERS coronaviruses (21).

Intensive testing and case-based interventions have so far formed the centerpiece of control efforts in some places, such as Singapore and Hong Kong (22). Many other countries are adopting measures termed “social distancing” or “physical distancing,” closing schools and workplaces and limiting the sizes of gatherings. The goal of these strategies is to reduce the peak intensity of the epidemic (“flatten the curve”) (22), reducing the risk of overwhelming health systems and buying time to develop treatments and vaccines. For social distancing to have reversed the epidemic in China, the effective reproduction number must have declined by at least 50-60%, assuming a baseline \( R_0 \) between 2 and 2.5 (22). Through intensive control measures, Shenzhen was able to reduce the effective reproduction number by an estimated 85% (23). However, it is unclear how well these declines in \( R_0 \) might generalize to other settings: recent data from Seattle suggests that the basic reproduction number has only declined to about 1.4, or by about 30-45% assuming a baseline \( R_0 \) between 2 and 2.5 (24). Furthermore, social distancing measures may need to last for months to effectively control transmission and mitigate the possibility of resurgence (25).

A key metric for the success of social distancing interventions is whether critical care capacities are exceeded. Modeling studies (26) and experience from the Wuhan outbreak (2) indicate that critical care capacities even in high-income countries can be exceeded many times over if distancing measures are not implemented quickly or strongly enough. To alleviate these problems, approaches to increase critical care capacity have included rapid construction or repurposing of hospital facilities and consideration of increased manufacturing and distribution of ventilators (27–30). Treatments that reduce the proportion of infections that lead to severe illness could have a similar effect of reducing burden on healthcare systems.

This paper identifies viral, environmental, and immunologic factors which in combination will determine the dynamics of SARS-CoV-2. We integrate our findings in a mathematical model to project potential scenarios for SARS-CoV-2 transmission through the pandemic and post-pandemic periods and identify key data still needed to determine which scenarios are likely to play out. Then, using the model, we assess the duration and intensity of social distancing measures that might be needed to maintain control of SARS-CoV-2 in the coming months under both existing and expanded critical care capacities.

**Transmission dynamics of HCoV-OC43 and HCoV-HKU1**

We used data from the United States to model betacoronavirus transmission in temperate regions and to project the possible dynamics of SARS-CoV-2 infection through the year 2025. We first assessed the role of seasonal variation, duration of immunity, and cross immunity on the transmissibility of HCoV-OC43 and HCoV-HKU1 in the US. We used the weekly percentage of positive laboratory tests for HCoV-OC43 and HCoV-HKU1 (31) multiplied by the weekly population-weighted proportion of physician visits due to influenza-like illness (ILI) (32, 33) to approximate historical betacoronavirus incidence in the US to within a scaling constant. This proxy is proportional to incidence under a set of assumptions described in the supplementary materials and methods. To quantify variation in transmission strength over time, we estimated the weekly effective reproduction number, defined as the average number of secondary infections caused by a single infected individual (34, 35). The effective reproduction numbers for each of the betacoronaviruses displayed a seasonal pattern, with annual peaks in the effective reproduction number slightly preceding those of the incidence curves (fig. S1). We limited our analysis to “in-season” estimates based on adequate samples, defined as week 40 through week 20 of the following year, roughly October to May. For both HCoV-OC43 and HCoV-HKU1, the effective reproduction number typically reached its peak between October and November and its trough between February and May. Over the five seasons included in our data (2014–2019), the median peak effective reproduction number was 1.85 (range: 1.61–2.21) for HCoV-HKU1 and 1.56 (range: 1.54–1.80) for HCoV-OC43 after removing outliers (five for HCoV-HKU1, zero for HCoV-OC43). Results were similar using various choices of incidence proxy and serial interval distributions (figs. S1 to S3).

To quantify the relative contribution of immunity versus seasonal forcing on the transmission dynamics of the betacoronaviruses, we adapted a regression model (36) that expressed the effective reproduction number for each strain (HKU1 and OC43) as the product of a baseline transmissibility constant (related to the basic reproduction number \( R_0 \)) and the proportion of the population susceptible at the start of each season), the depletion of susceptibles due to infection with the same strain, the depletion of susceptibles due to infection with the other strain, and a spline to capture further
unexplained seasonal variation in transmission strength (seasonal forcing). These covariates were able to explain most of the observed variability in the effective reproduction numbers (adjusted R²: 74.3%). The estimated multiplicative effects of each of these covariates on the weekly reproduction number are depicted in Fig. 1. As expected, depletion of susceptibles for each strain was negatively correlated with transmissibility of that strain. Depletion of susceptibles for each strain was also negatively correlated with the reproduction number of the other betacoronavirus strain, providing evidence of cross-immunity. Per incidence proxy unit, the effect of the cross-immunizing strain was always less than the effect of the strain itself (table S1), but the overall impact of cross-immunity on the reproduction number could still be substantial if the cross-immunizing strain had a large outbreak (e.g., HCoV-OC43 in 2014-15 and 2016-17). The ratio of cross-immunization to self-immunization effects was larger for HCoV-HKU1 than for HCoV-OC43, suggesting that HCoV-OC43 confers stronger cross-immunity. Seasonal forcing appears to drive the rise in transmissibility at the start of the season (late October through early December), while depletion of susceptibles plays a comparatively larger role in the decline in transmissibility toward the end of the season. The strain-season coefficients were fairly consistent across seasons for each strain and lacked a clear correlation with incidence in prior seasons, consistent with experimental results showing substantial waning of immunity within a year (15).

We integrated these findings into a two-strain ordinary differential equation (ODE) susceptible-exposed-infectious-recovered-susceptible (SEIRS) compartmental model to describe the transmission dynamics of HCoV-OC43 and HCoV-HKU1 (fig. S4). The model provided a good fit to both the weekly incidence proxies for HCoV-OC43 and HCoV-HKU1 and to the estimated weekly effective reproduction numbers (Fig. 2). According to the best-fit model parameters, the R₀ for HCoV-OC43 and HCoV-HKU1 varies between 1.7 in the summer and 2.2 in the winter and peaks in the 2nd week of January, consistent with the seasonal spline estimated from the data. Also in agreement with the findings of the regression model, the duration of immunity for both strains in the best-fit SEIRS model is about 45 weeks, and each strain induces cross-immunity against the other, though the cross-immunity that HCoV-OC43 infection induces against HCoV-HKU1 is stronger than the reverse.

**Simulating the transmission of SARS-CoV-2**

Next, we incorporated a third betacoronavirus into the dynamic transmission model to represent SARS-CoV-2. We assumed a latent period of 4.6 days (26, 37–39) and an infectious period of 5 days, informed by the best-fit values for the other betacoronaviruses (table S8). We allowed the cross immunities, duration of immunity, maximum R₀, and degree of seasonal variation in R₀ to vary. We assumed an establishment time of sustained transmission on 11 March 2020, when the World Health Organization declared the SARS-CoV-2 outbreak a pandemic (40) and we varied the establishment time in a sensitivity analysis (fig. S7). For a representative set of parameter values, we measured annual SARS-CoV-2 infections (tables S2 to S4 and fig. S7) and the peak annual SARS-CoV-2 prevalence (tables S5 to S7 and fig. S7) through 2025. We summarized the post-pandemic SARS-CoV-2 dynamics into the categories of annual outbreaks, biennial outbreaks, sporadic outbreaks, or virtual elimination (tables S2 to S7). Overall, shorter durations of immunity and smaller degrees of cross-immunity from the other betacoronaviruses were associated with greater total incidence of infection due to SARS-CoV-2, and annual establishments and smaller seasonal fluctuations in transmissibility were associated with larger pandemic peak sizes. Model simulations demonstrated the following key points.

**SARS-CoV-2 can proliferate at any time of year**

In all modeled scenarios, SARS-CoV-2 was capable of producing a substantial outbreak regardless of establishment time. Winter/spring establishments favored outbreaks with lower peaks, while autumn/winter establishments led to more acute outbreaks (tables S2 to S4 and fig. S7). The five-year cumulative incidence proxies were comparable for all establishment times (tables S2 to S4).

**If immunity to SARS-CoV-2 is not permanent, it will likely enter into regular circulation**

Much like pandemic influenza, many scenarios lead to SARS-CoV-2 entering into long-term circulation alongside the other human betacoronaviruses (e.g., Fig. 3, A and B), possibly in annual, biennial, or sporadic patterns over the next five years (tables S2 to S4). Short-term immunity (on the order of 40 weeks, similar to HCoV-OC43 and HCoV-HKU1) favors the establishment of annual SARS-CoV-2 outbreaks, while longer-term immunity (two years) favors biennial outbreaks.

**High seasonal variation in transmission leads to smaller peak incidence during the initial pandemic wave but larger recurrent wintertime outbreaks**

The amount of seasonal variation in SARS-CoV-2 transmission could differ between geographic locations, as for influenza (12). The R₀ for influenza in New York declines in the summer by about 40%, while in Florida the decline is closer to 20%, which aligns with the estimated decline in R₀ for HCoV-OC43 and HCoV-HKU1 (table S8). A 40% summertime decline in R₀ would reduce the unmitigated peak incidence of the initial SARS-CoV-2 pandemic wave. However, stronger seasonal forcing leads to a greater accumulation of susceptible individuals during periods of low transmission in the
summer, leading to recurrent outbreaks with higher peaks in the post-pandemic period (Fig. 3C).

**If immunity to SARS-CoV-2 is permanent, the virus could disappear for five or more years after causing a major outbreak**

Long-term immunity consistently led to effective elimination of SARS-CoV-2 and lower overall incidence of infection. If SARS-CoV-2 induces cross immunity against HCoV-OC43 and HCoV-HKU1, the incidence of all betacoronaviruses could decline and even virtually disappear (Fig. 3D). The virtual elimination of HCoV-OC43 and HCoV-HKU1 would be possible if SARS-CoV-2 induced 70% cross immunity against them, which is the same estimated level of cross-immunity that HCoV-OC43 induces against HCoV-HKU1.

**Low levels of cross immunity from the other betacoronaviruses against SARS-CoV-2 could make SARS-CoV-2 appear to die out, only to resurge after a few years**

Even if SARS-CoV-2 immunity only lasts for two years, mild (30%) cross-immunity from HCoV-OC43 and HCoV-HKU1 could effectively eliminate the transmission of SARS-CoV-2 for up to three years before a resurgence in 2024, as long as SARS-CoV-2 does not fully die out (Fig. 3E).

To illustrate these scenarios (Fig. 3), we used a maximum wintertime $R_0$ of 2.2, informed by the estimated $R_0$ for HCoV-OC43 and HCoV-HKU1 (table S8). This is a low but plausible estimate of the basic reproduction number for SARS-CoV-2 (41). Increasing wintertime $R_0$ to 2.6 leads to more intense outbreaks, but the qualitative range of scenarios remains similar (fig. S8).

**Assessing intervention scenarios during the initial pandemic wave**

Regardless of the post-pandemic transmission dynamics of SARS-CoV-2, urgent measures are required to address the ongoing epidemic. Pharmaceutical treatments and vaccines may require months to years to develop and test, leaving non-pharmaceutical interventions (NPIs) as the only immediate means of curbing SARS-CoV-2 transmission. Social distancing measures have been adopted in many countries with widespread SARS-CoV-2 transmission. The necessary duration and intensity of these measures has yet to be characterized. To address this, we adapted the SEIRS transmission model (fig. S9) to capture moderate/mild/asymptomatic infections (95.6% of infections), infections that lead to hospitalization but not critical care (3.08% of infections), and infections that require critical care (1.32% of infections) (26).

We assumed the worst-case scenario of no cross-immunity from HCoV-OC43 and HCoV-HKU1 against SARS-CoV-2, which makes the SARS-CoV-2 model unaffected by the transmission dynamics of those viruses. Informed by the transmission model fits, we assumed a latent period of 4.6 days and an infectious period of 5 days, in agreement with estimates from other studies (26). The mean duration of non-critical hospital stay was 8 days for those not requiring critical care and 6 days for those requiring critical care, and the mean duration of critical care was 10 days (26). We varied the peak (wintertime) $R_0$ between 2.2 and 2.6 and allowed the summertime $R_0$ to vary between 60% (i.e., relatively strong seasonality) and 100% (i.e., no seasonality) of the wintertime $R_0$, guided by the inferred seasonal forcing for HCoV-OC43 and HCoV-HKU1 (table S8).

We used the open critical care capacity of the United States, 0.89 free beds per 10,000 adults, as a benchmark for critical care demand (2). We simulated epidemic trajectories based on an epidemic establishment time of 11 March 2020. We simulated social distancing by reducing $R_0$ by a fixed proportion, which ranged between 0 and 60%. We assessed ‘one-time’ social distancing interventions, for which $R_0$ was reduced by up to 60% for a fixed duration of time (up to 20 weeks) or indefinitely starting two weeks after epidemic establishment. We also assessed intermittent social distancing measures, for which social distancing was turned ‘on’ when the prevalence of infection rose above a threshold and ‘off’ when it fell below a second, lower threshold, with the goal of keeping the number of critical care patients below 0.89 per 10,000 adults. An ‘on’ threshold of 35 cases per 10,000 people achieved this goal in both the seasonal and non-seasonal cases with wintertime $R_0 = 2.2$. We chose 5 cases per 10,000 adults as the ‘off’ threshold. These thresholds were chosen to qualitatively illustrate the intermittent intervention scenario; in practice, the thresholds will need to be tuned to local epidemic dynamics and hospital capacities. We performed a sensitivity analysis around these threshold values (figs. S10 and S11) to assess how they affected the duration and frequency of the interventions. We also implemented a model with extra compartments for the latent period, infectious period, and each hospitalization period so that the waiting times in these states were gamma-distributed instead of exponentially distributed (see the supplementary materials and methods and figs. S16 and S17). Finally, we assessed the impact of doubling critical care capacity (and the associated on/off thresholds) on the frequency and overall duration of the social distancing measures.

We evaluated the impact of one-time social distancing efforts of varying effectiveness and duration on the peak and timing of the epidemic with and without seasonal forcing. When transmission was not subject to seasonal forcing, one-time social distancing measures reduced the epidemic peak size (Fig. 4 and fig. S12). Under all scenarios, there was a resurgence of infection when the simulated social distancing measures were lifted. However, longer and more stringent temporary social distancing did not always correlate with
greater reductions in epidemic peak size. In the case of a 20-week period of social distancing with 60% reduction in $R_0$, for example (Fig. 4D), the resurgence peak size was nearly the same as the peak size of the uncontrolled epidemic: the social distancing was so effective that virtually no population immunity was built. The greatest reductions in peak size come from social distancing intensity and duration that divide cases approximately equally between peaks (42).

For simulations with seasonal forcing, the post-intervention resurgence peak could exceed the size of the unconstrained epidemic (Fig. 5 and fig. S13), both in terms of peak prevalence and in terms of total number infected. Strong social distancing maintained a high proportion of susceptible individuals in the population, leading to an intense epidemic when $R_0$ rises in the late autumn and winter. None of the one-time interventions was effective in maintaining the prevalence of critical cases below the critical care capacity.

Intermittent social distancing could prevent critical care capacity from being exceeded (Fig. 6 and fig. S14). Due to the natural history of infection, there is an approximately 3-week lag between the start of social distancing and the peak critical care demand. When transmission is seasonally forced, summertime social distancing can be less frequent than when $R_0$ remains constant at its maximal wintertime value throughout the year. The length of time between distancing measures increases as the epidemic continues, as the accumulation of immunity in the population slows the resurgence of infection. Under current critical care capacities, however, the overall duration of the SARS-CoV-2 epidemic could last into 2022, requiring social distancing measures to be in place between 25% (for wintertime $R_0 = 2$ and seasonality; fig. S11A) and 75% (for wintertime $R_0 = 2.6$ and no seasonality; fig. S9C) of that time. When the latent, infectious, and hospitalization periods are gamma-distributed, incidence rises more quickly, requiring a lower threshold for implementing distancing measures (25 cases per 10,000 individuals for $R_0 = 2.2$ in our model) and more frequent interventions (fig. S16).

Increasing critical care capacity allowed population immunity to be accumulated more rapidly, reducing the overall duration of the epidemic and the total length of social distancing measures (Fig. 6, C and D). While the frequency and duration of the social distancing measures was similar between the scenarios with current and expanded critical care capacity, the epidemic concluded by July 2022 and social distancing measures could be fully relaxed by early-to mid-2021, depending again on the degree of seasonal forcing of transmission (Fig. 6, C and D). Introducing a hypothetical protocol that halved the proportion of infections that required hospitalization had a similar effect as doubling critical care capacity (fig. S15).

Discussion

Here, we examined a range of likely SARS-CoV-2 transmission scenarios through 2025 and assessed non-pharmaceutical interventions that could mitigate the intensity of the current outbreak. If immunity to SARS-CoV-2 wanes in the same manner as related coronaviruses, recurrent wintertime outbreaks are likely to occur in coming years. The total incidence of SARS-CoV-2 through 2025 will depend crucially on this duration of immunity and, to a lesser degree, on the amount of cross immunity that exists between HCoVs OC43/HKU1 and SARS-CoV-2. The intensity of the initial pandemic wave will depend fundamentally on the basic reproduction number at the time of epidemic establishment: if establishment occurs in the autumn when the reproduction number is rising, which could occur in countries that maintain epidemic control by contact tracing and quarantine through the summer, or if SARS-CoV-2 is not subject to the same summertime decline in transmissibility as HCoV-OC43 and HCoV-HKU1, then a high peak prevalence of infection is likely. One-time social distancing efforts may push the SARS-CoV-2 epidemic peak into the autumn, potentially exacerbating the load on critical care resources if there is increased wintertime transmissibility. Intermittent social distancing might maintain critical care demand within current thresholds, but widespread surveillance will be required to time the distancing measures correctly and avoid overshooting critical care capacity. New therapeutics, vaccines, or other interventions such as aggressive contact tracing and quarantine – impractical now in many places but more practical once case numbers have been reduced and testing scaled up (43) – could alleviate the need for stringent social distancing to maintain control of the epidemic. In the absence of such interventions, surveillance and intermittent distancing (or sustained distancing if it is highly effective) may need to be maintained into 2022, which would present a substantial social and economic burden. To shorten the SARS-CoV-2 epidemic and ensure adequate care for the critically ill, increasing critical care capacity and developing additional interventions are urgent priorities. Meanwhile, serological testing is required to understand the extent and duration of immunity to SARS-CoV-2, which will help determine the post-pandemic dynamics of the virus. Sustained, widespread surveillance will be needed both in the short term to effectively implement intermittent social distancing measures and in the long term to assess the possibility of resurgences of SARS-CoV-2 infection, which could occur as late as 2025 even after a prolonged period of apparent elimination.

Our observations are consistent with other predictions of how SARS-CoV-2 transmission might unfold and with assessments of the mitigation efforts that might be needed to curb the current outbreak. A modelling study using data from Sweden found that seasonal establishment of SARS-CoV-2 transmission is likely in the post-pandemic period (17). Observational and modelling studies (2, 26) that find that early implementation of strong social distancing is essential.
for controlling the spread of SARS-CoV-2 and that, in the absence of the development of new therapies or preventative measures, such as aggressive case finding and quarantining (27), intermittent distancing measures may be the only way to avoid overwhelming critical care capacity while building population immunity. The observation that strong, temporary social distancing can lead to especially large resurgences agrees with data from the 1918 influenza pandemic in the United States (44), in which the size of the autumn 1918 peak of infection was inversely associated with that of a subsequent winter peak after interventions were no longer in place.

Our study was subject to a variety of limitations. Only five seasons of observational data on coronaviruses were available, though the incidence patterns resemble those from 10 years of data from a hospital in Sweden (11). We assumed that the spline coefficients were constant across all seasons though seasonal forcing likely differed from year to year based on underlying drivers. To keep the transmission model from becoming unreasonably complex, we assumed that there was no difference in the seasonal forcing, per-case force of infection, latent period, or infectious period across betacoronaviruses. However, our estimates for these values lie within the ranges of estimates from the literature. Although disease dynamics may differ by age, we did not have sufficient data to parameterize an age-structured model. We also did not directly model any effect from the opening of schools, which could lead to an additional boost in transmission strength in the early autumn (45). The transmission model is deterministic, so it cannot capture the possibility of SARS-CoV-2 extinction. It also does not incorporate geographic structure, so the possibility of spatially heterogeneous transmission cannot be assessed. The construction of spatially explicit models will become more feasible as more data on SARS-CoV-2 incidence becomes available; these will help determine whether there are differences in seasonal forcing between geographic locations, as for influenza (12), and will also help to assess the possibility of epidemic extinction while accounting for re-introductions. The timing and strength of post-pandemic outbreaks may also depend on stochastic introductions from abroad, which can be assessed using more complex, global models.

We used percent test-positive multiplied by percent ILI to approximate coronavirus incidence up to a proportional constant; results were similar when using the raw number of positive tests and the raw percent positive as incidence proxies (fig. S1). While the percent test-positive multiplied by percent ILI has been shown to be one of the best available proxies for influenza incidence (32), the conversion between this measure and true incidence of coronavirus infections is unclear, and so we do not make precise estimates of the overall coronavirus incidence. This conversion will undoubtedly depend on the particular population for which these estimates are being made. In a recent study, an estimated 4% of individuals with coronavirus sought medical care, and only a fraction of these were tested (46). In addition, the method that we adopted to estimate the effective reproduction number depends on the serial interval distribution, which has not been well-studied for commonly circulating human coronaviruses; we used the best-available evidence from SARS-CoV-1, the most closely related coronavirus to SARS-CoV-2.

Our findings generalize only to temperate regions, comprising 60% of the world's population (47), and the size and intensity of outbreaks could be further modulated by differences in average interpersonal contact rates by location and the timing and effectiveness of non-pharmaceutical and pharmaceutical interventions. The transmission dynamics of respiratory illnesses in tropical regions can be much more complex. However, we expect that if post-pandemic transmission of SARS-CoV-2 does take hold in temperate regions, there will also be continued transmission in tropical regions seeded by the seasonal outbreaks to the north and south. With such reseeding, long-term disappearance of any strain becomes less likely (48), but according to our model the effective reproductive number of SARS-CoV-2 remains below 1 during most of each period when that strain disappears, meaning that reseeding would shorten these disappearances only modestly.

Our findings indicate key data required to know how the current SARS-CoV-2 outbreak will unfold. Most crucially, serological studies could indicate the extent of population immunity, whether immunity wanes, and at what rate. In our model, this rate is the key modulator of the total SARS-CoV-2 incidence in the coming years. While long-lasting immunity would lead to lower overall incidence of infection, it would also complicate vaccine efficacy trials by contributing to low case numbers when those trials are conducted, as occurred with Zika virus (49). In our assessment of control measures in the initial pandemic period, we assumed that SARS-CoV-2 infection induces immunity that lasts for at least two years, but social distancing measures may need to be extended if SARS-CoV-2 immunity wanes more rapidly. In addition, if serological data reveals the existence of many undocumented asymptomatic infections that lead to immunity (50), less social distancing may be required. Serology could also indicate whether cross-immunity exists between SARS-CoV-2, HCoV-OC43, and HCoV-HKU1, which could affect the post-pandemic transmission of SARS-CoV-2. We anticipate that such cross-immunity would lessen the intensity of SARS-CoV-2 outbreaks, though some speculate that antibody-dependent enhancement (ADE) induced by prior coronavirus infection may increase susceptibility to SARS-CoV-2 and exacerbate the severity of infection (51, 52). At present, there is limited data describing ADE between coronaviruses, but if it does exist it
may promote the cocirculation of betacoronavirus strains.

To implement intermittent social distancing, it will be necessary to carry out widespread viral testing for surveillance to monitor when the prevalence thresholds that trigger the beginning or end of distancing have been crossed. Without such surveillance, critical care bed availability might be used as a proxy for prevalence, but this metric is far from optimal since the lag between distancing and peak critical care demand could lead to frequent overrunning of critical care resources. Critical care resources are also at greater risk of being overrun if the infectious, latent, and hospitalized periods follow peaked distributions (e.g., gamma versus exponential). Measuring the distributions of these times, and not just their means, will help to set more effective thresholds for distancing interventions. Under some circumstances, intense social distancing may be able to reduce the prevalence of COVID-19 enough to warrant a shift in strategy to contact tracing and containment efforts, as has occurred in many parts of China (21, 23, 53). Still, countries that have achieved this level of control of the outbreak should prepare for the possibility of substantial resurgences of infection and a return to social distancing measures, especially if seasonal forcing contributes to a rise in transmissibility in the winter. Moreover, a winter peak for COVID-19 would coincide with peak influenza incidence (54), further straining health care systems.

Treatments or vaccines for SARS-CoV-2 would reduce the duration and intensity of social distancing required to maintain control of the epidemic. Treatments could reduce the proportion of infections that require critical care and could reduce the duration of infectiousness, which would both directly and indirectly (through a reduction in $R_0$) reduce the demand for critical care resources. A vaccine would accelerate the accumulation of immunity in the population, reducing the overall length of the epidemic and averting infections that might have resulted in a need for critical care. Furthermore, if there have been many undocumented immunizing infections, the herd immunity threshold may be reached sooner than our models suggest. Still, SARS-CoV-2 has demonstrated an ability to challenge robust healthcare systems, and the development and widespread adoption of pharmaceutical interventions will take months at best, so a period of sustained or intermittent social distancing will almost certainly be necessary.

In summary, the total incidence of COVID-19 illness over the next five years will depend critically upon whether or not it enters into regular circulation after the initial pandemic wave, which in turn depends primarily upon the duration of immunity that SARS-CoV-2 infection imparts. The intensity and timing of pandemic and post-pandemic outbreaks will depend on the time of year when widespread SARS-CoV-2 infection becomes established and, to a lesser degree, upon the magnitude of seasonal variation in transmissibility and the level of cross-immunity that exists between the betacoronaviruses. Social distancing strategies could reduce the extent to which SARS-CoV-2 infections strain health care systems. Highly-effective distancing could reduce SARS-CoV-2 incidence enough to make a strategy based on contact tracing and quarantine feasible, as in South Korea and Singapore. Less effective one-time distancing efforts may result in a prolonged single-peak epidemic, with the extent of strain on the healthcare system and the required duration of distancing depending on the effectiveness. Intermittent distancing may be required into 2022 unless critical care capacity is increased substantially or a treatment or vaccine becomes available. The authors are aware that prolonged distancing, even if intermittent, is likely to have profoundly negative economic, social, and educational consequences. Our goal in modeling such policies is not to endorse them but to identify likely trajectories of the epidemic under alternative approaches, identify complementary interventions such as expanding ICU capacity and identifying treatments to reduce ICU demand, and to spur innovative ideas (55) to expand the list of options to bring the pandemic under long-term control. Our model presents a variety of scenarios intended to anticipate possible SARS-CoV-2 transmission dynamics under specific assumptions. We do not take a position on the advisability of these scenarios given the economic burden that sustained distancing may impose, but we note the potentially catastrophic burden on the healthcare system that is predicted if distancing is poorly effective and/or not sustained for long enough. The model will have to be tailored to local conditions and updated as more accurate data become available. Longitudinal serological studies are urgently required to determine the extent and duration of immunity to SARS-CoV-2, and epidemiological surveillance should be maintained in the coming years to anticipate the possibility of resurgence.

REFERENCES AND NOTES
3. “‘Not a wave, a tsunami.’ Italy hospitals at virus limit,” AP NEWS, 13 March 2020; https://apnews.com/9d4573d1b5d5-d8f2b7b6e43e539f669.
doi:10.1016/S0140-6736(13)60585-3 Medicine
doi:10.1098/rspb.2006.0015 Medianline


doi:10.1038/srep11013 Medianline


C. Tedjianto, c2-d2/CoV-seasonality: First release, Zenodo (2020); https://zenodo.org/record/3726035#.XcG6s_7hwkk.

S. Kissler, nCoV_introduction, Version 4.0, Zenodo (2020); https://zenodo.org/record/4745557#.XpC7Y_7hwkw


M. D. McKay, R. J. Beckman, W. J. Conover. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code.


ACKNOWLEDGMENTS

We thank Marie Killery and Amber Haynes for their helpful comments on early versions of this manuscript. We thank M.W. Shelley for advice on structuring the manuscript and Benedicte Gnanngon for helpful discussions on viral immunity dynamics. Funding: CT was supported by T22AI07535 from the National Institute of Allergy and Infectious Diseases. The work was also supported by the Morris-Singer Fund for the Center for Communicable Disease Dynamics at the Harvard T.H. Chan School of Public Health. Y.G. was funded by the NIH, Doris Duke Charitable Foundation, the Wellcome Trust, and internal awards through Harvard University. Author contributions: SMK conceived of the study, conducted the analysis, and wrote the manuscript. CT conceived of the study, conducted the analysis, and wrote the manuscript. EG assisted with the analysis and edited the manuscript. YHG conceived of the study, edited the manuscript and oversaw the work. ML conceived of the study, edited the manuscript and oversaw the work. Competing interests: The authors declare no competing interests. Data and materials availability: UNInet data are publicly available through the FluView Interactive website (33). Regression (56) and transmission model (57) code are available in online repositories. This work is licensed under a Creative Commons Attribution 4.0 International (CC BY 4.0) license, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. To view a copy of this license, visit https://creativecommons.org/licenses/by/4.0/. This license does not apply to figures/photos/artwork or other content included in the article that is credited to a third party; obtain authorization from the rights holder before using such material.

SUPPLEMENTARY MATERIALS

science-sciencemag.org/cgi/content/full/science.abb5793/DC1

Materials and Methods

Figs. S1 to S17

Tables S1 to S8

References (56–62)

MDR Reproducibility Checklist

4 March 2020; accepted 9 April 2020

Published online 14 April 2020

10.1126/science.abb5793

Downloaded from http://science.sciencemag.org/ on August 5, 2021
Fig. 1. Effects of depletion of susceptibles and seasonality on the effective reproduction number by strain and season. Estimated multiplicative effects of HCoV-HKU1 incidence (red), HCoV-OC43 incidence (blue), and seasonal forcing (gold) on weekly effective reproduction numbers of HCoV-HKU1 (top panels) and HCoV-OC43 (bottom), with 95% confidence intervals. The black dot (with 95% confidence interval) plotted at the start of each season is the estimated coefficient for that strain and season compared to the 2014-15 HCoV-HKU1 season. The seasonal forcing spline is set to 1 at the first week of the season (no intercept). On the x-axis, the first “week in season” corresponds to epidemiological week 40.
Fig. 2. Transmission model fits for HCoV-OC43 and HCoV-HKU1. (A) Weekly percent positive laboratory tests multiplied by percent influenza-like illness (ILI) for the human betacoronaviruses HCoV-OC43 (blue) and HCoV-HKU1 (red) in the United States between 5 July 2014 and 29 June 2019 (solid lines) with simulated output from the best-fit SEIRS transmission model (dashed lines). (B and C) Weekly effective reproduction numbers ($R_e$) estimated using the Wallinga-Teunis method (points) and simulated $R_e$ from the best-fit SEIRS transmission model (line) for HCoVs OC43 and HKU1. The opacity of each point is determined by the relative percent ILI multiplied by percent positive laboratory tests in that week relative to the maximum percent ILI multiplied by percent positive laboratory tests for that strain across the study period, which reflects uncertainty in the $R_e$ estimate; estimates are more certain (darker points) in weeks with higher incidence.
Fig. 3. Invasion scenarios for SARS-CoV-2 in temperate regions. These plots depict the prevalence of SARS-CoV-2 (black, cases per 1,000 people), HCoV-OC43 (blue, % positive multiplied by % ILI), and HCoV-HKU1 (red, % positive multiplied by % ILI) for a representative set of possible pandemic and post-pandemic scenarios. The scenarios were obtained by varying the cross immunity between SARS-CoV-2 and HCoVs OC43/HKU1 (χ_{3X}) and vice-versa (χ_{X3}), the duration of SARS-CoV-2 immunity (1/α_{3}), and the seasonal variation in R_{0} (f), assuming an epidemic establishment time of 11 March 2020 (depicted as a vertical grey bar). Parameter values used to generate each plot are listed below; all other parameters were held at the values listed in table S8. (A) A short duration (1/α_{3} = 40 weeks) of SARS-CoV-2 immunity could yield annual SARS-CoV-2 outbreaks. (B) Longer-term SARS-CoV-2 immunity (1/α_{3} = 104 weeks) could yield biennial outbreaks, possibly with smaller outbreaks in the intervening years. (C) Higher seasonal variation in transmission (f = 0.4) would reduce the peak size of the invasion wave, but could lead to more severe wintertime outbreaks thereafter [compare with (B)]. (D) Long-term immunity (1/α_{3} = infinity) to SARS-CoV-2 could lead to elimination of the virus. (E) However, a resurgence of SARS-CoV-2 could occur as late as 2024 after a period of apparent elimination if the duration of immunity is intermediate (1/α_{3} = 104 weeks) and if HCoVs OC43/HKU1 impart intermediate cross immunity against SARS-CoV-2 (χ_{3X} = 0.3). (A) χ_{X3} = 0.3, χ_{X3} = 0, 1/α_{3} = 40 weeks, f = 0.2. (B) χ_{3X} = 0.7, χ_{X3} = 0, 1/α_{3} = 104 weeks, f = 0.2. (C) χ_{3X} = 0.7, χ_{X3} = 0, 1/α_{3} = 104 weeks, f = 0.4. (D) χ_{3X} = 0.7, χ_{X3} = 0, 1/α_{3} = infinity, f = 0.2. (E) χ_{3X} = 0.3, χ_{X3} = 0.3, 1/α_{3} = 104 weeks, f = 0.4.
Fig. 4. One-time social distancing scenarios in the absence of seasonality. (A to E) Simulated prevalence of COVID-19 infections (solid) and critical COVID-19 cases (dashed) following establishment on 11 March 2020 with a period of social distancing (shaded blue region) instated two weeks later, with the duration of social distancing lasting (A) four weeks, (B) eight weeks, (C) twelve weeks, (D) twenty weeks, and (E) indefinitely. There is no seasonal forcing; $R_0$ was held constant at 2.2 (see fig. S12 for $R_0 = 2.6$). The effectiveness of social distancing varied from none to a 60% reduction in $R_0$. Cumulative infection sizes are depicted beside each prevalence plot (F to J) with the herd immunity threshold (horizontal black bar). Of the temporary distancing scenarios, long-term (20-week), moderately effective (20%-40%) social distancing yields the smallest overall peak and total outbreak size.
Fig. 5. One-time social distancing scenarios with seasonal transmission. (A to E) Simulated prevalence, assuming strong seasonal forcing (wintertime $R_0 = 2.2$, summertime $R_0 = 1.3$, or 40% decline), of COVID-19 infections (solid) and critical COVID-19 cases (dashed) following establishment on 11 March 2020 with a period of social distancing (shaded blue region) instated two weeks later, with the duration of social distancing lasting (A) four weeks, (B) eight weeks, (C) twelve weeks, (D) twenty weeks, and (E) indefinitely (see fig. S13 for a scenario with wintertime $R_0 = 2.6$). The effectiveness of social distancing varied from none to a 60% reduction in $R_0$. Cumulative infection sizes are depicted beside each prevalence plot (F to J) with the herd immunity threshold (horizontal black bar). Preventing widespread infection during the summer can flatten and prolong the epidemic but can also lead to a high density of susceptible individuals who could become infected in an intense autumn wave.
Fig. 6. Intermittent social distancing scenarios with current and expanded critical care capacity. SARS-CoV-2 prevalence (black curves) and critical cases (red curves) under intermittent social distancing (shaded blue regions) without seasonal forcing (A and C) and with seasonal forcing (B and D). Distancing yields a 60% reduction in \( R_0 \). Critical care capacity is depicted by the solid horizontal black bars, and the on/off thresholds for social distancing are depicted by the dashed horizontal lines. (A) and (B) are the scenarios with current US critical care capacity and (C) and (D) are the scenarios with double the current critical care capacity. The maximal wintertime \( R_0 \) is 2.2 and for the seasonal scenarios the summertime \( R_0 \) is 1.3 (40% decline). Prevalence is in black and critical care cases are in red. To the right of each main plot (E to H), the proportion immune over time is depicted in green with the herd immunity threshold (horizontal black bar).
Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period
Stephen M. Kissler, Christine Tedijanto, Edward Goldstein, Yonatan H. Grad and Marc Lipsitch

published online April 14, 2020 originally published online April 14, 2020