

Country-specific Optimization Strategy for Testing Through Contact Tracing Can Help Maintain a Low Reproduction Number (R_0) During Unlock

Uddipan Sarma

Vantage Research

Bhaswar Ghosh (✉ bhaswar.ghosh@iiit.ac.in)

International Institute of Information Technology

Research Article

Keywords: COVID19 pandemics, xtra burden on quarantine, testing capacity, unlock extents

Posted Date: May 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-523250/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Country-specific optimization strategy for testing through contact tracing can help maintain a low reproduction number (R_0) during unlock

Uddipan Sarma^{1,#}, Bhaswar Ghosh^{2,#}

¹Vantage Research, Sivasamy St, CIT Colony, Mylapore, Chennai, Tamil Nadu 600004. India.

²Center for Computational Natural Sciences, International Institute of Information Technology, Hyderabad 500032, India

Joint Correspondence :

Uddipan Sarma : uddipans@gmail.com

Bhaswar Ghosh : bhaswar.ghosh@iiit.ac.in

Abstract

In response to the COVID19 pandemics, many countries have implemented lockdowns in multiple phases to ensure social distancing and quarantining of the infected subjects as a first step to contain the infection spread. Subsequent unlocks to reopen the economies started next waves and imposed extra burden on quarantine to keep the reproduction number ($R_0 < 1$). Even with initial strict lockdowns and recent launching of vaccination programs, many countries are still struggling to contain the infection which suggests that revisiting the mechanism of lockdown-unlock implementation and simultaneous underpinning of the potential sources diluting the effort of such lockdowns could help better contain the spread of infection. Here, building epidemic models and analyzing the metadata of 50 countries, we first found that the estimated values of R_0 , adjusted w.r.t the distribution of medical facilities and virus clades, correlates strongly with the testing rates across countries. However, testing capacity of a country is limited by its medical resources, hence, as we demonstrate, optimizing a cost-benefit trade-off between testing rate and unlocking extents implemented in a country specific manner can help in devising the strategies of unlocking the economy. Our study delineates a strategy to optimize this trade-off by utilizing country specific infection spread parameters estimated in the epidemic models and implementing them in a stochastic agent based contact tracing models. The analysis provides a quantitative estimate of testing rates required to maintain a low R_0 for different extents of unlock. We further found that a small fraction of superspreaders can drastically increase the number of infected individuals even during lockdowns, primarily due to a switch-like response stemming from the implicit systems-level positive feedback loop driving the spread of infection. Our model suggests that with a country specific optimal combination of unlock extents and testing rates, $R_0 < 1$ can be stabilized during

a pandemic like COVID19. To harness the benefit of improved testing rates and minimize the infection spread, strict social distancing norms to [restrict the movement of superspreaders is necessary, such that onset of the positive feedback loop mediated exponential infection spread can be avoided.](#)

Introduction

Declaration of the coronavirus pandemic by WHO severely overhauled global economic and social endeavors¹. In response to initial upsurge of the infection spread^{3,4}, many European and east asian countries were able to contain the first wave successfully by imposing strong mitigation measures through nationwide lock-down coupled with rigorous testing and quarantine strategies. However, even with [strict lock-downs many countries struggled to contain the growth of infection². During the vaccine preparation phase⁵, isolating the infected population by aggressive testing and maintaining strong social distancing measures are adopted globally as the two most effective ways to contain the infection spread and reduce the fatalities ^{6,7}. Despite recent launching of several vaccine programs worldwide, complete immunization can take several months to a couple of years to cover the complete cross-section of susceptible age groups in all countries. Also there are indications that vaccination alone may not be sufficient to contain COVID19⁸. This inevitably rearticulates the importance of occasional lockdown and quarantine strategies in the immediate future.](#)

Despite benefits like saving numerous lives, long lockdowns imposed at the beginning severely disrupted the livelihood of millions of people, especially those of low and medium income countries⁹⁻¹² calling for immediate action plans to open up the economic activities. The unlocking measures subsequently led to an increase in infection rates which initiated the next wave of infection in multiple countries that were largely successful in their attempt to contain the first wave. For instance, a second and subsequent waves of infection are being encountered in multiple regions of Europe or the US. [Thus the lockdowns were not successful to the extent desired and there could be scope of improvement in the strategies of testing and quarantine in order to maintain the basic economic activities while also ensuring a minimal spread of infection during the unlocks and lockdowns. In the global context, the number of testing conducted per day in a country is limited by the resource capacity¹³ and the testing capacity is steadily rising in most of the countries ¹⁴⁻¹⁶. Further, the allowed extent of unlocking in a country can be a function of testing capacity as the daily testing captures the increase/decrease in infections per day. However it is not well established how testing facilities and extent of unlock are connected and a country-specific cost-benefit trade-off approach can potentially aim at devising quantitative measures of unlocking in view of the maximum resource capacity for testing such that](#)

infection rate during the unlocking is minimized.

Here, building a dynamic epidemic model¹⁷⁻²⁰ for the spread of coronavirus and quantitatively calibrating the time series data¹ for confirmed, recovered and dead population for 50 different countries with various stages of infection, we first made an estimate of parameters like incubation time, transmission rate, quarantine, recovery and death rates. We introduced dynamically adjustable lockdowns in our model to observe the effect of social distancing and also estimated the effectiveness of implementation of lock down in individual countries by fitting a lockdown parameter. Using the parameters from the epidemic model we next analytically determined the basic reproduction number (R_0) and developed a model to implement the country specific cost-benefit trade-off between different extent of partial unlocks and quarantine rates. The analysis shows how extent of partial unlock and quarantine rates can be optimally combined to maintain $R_0 < 1$. We also connected the extent of unlock with the frequency of the periodic unlock²¹ time. Next as the dynamic epidemic models build with ordinary differential equation (ODE) can not explicitly capture the relation between R_0 and testing rates, we developed a country specific agent based stochastic version of the epidemic model which incorporates testing through contact tracing of individual infected agents.. Using this hybrid modeling approach we first estimated the testing rate required to maintain $R_0 < 1$ and further demonstrate that a very small fraction of superspreaders²² can dramatically spread the infection even during lockdowns. Indeed, many countries report continuous rise in infection during lockdown. In fact, recent study on two Indian states showed a small fraction of superspreaders were responsible for transmitting 80% of infection during lockdown²². Analysis of the country specific models show a switch-like spread of infection can occur as a function of transmission and quarantine rates and this switch-like response stems from the inherent systems-level positive feedback loop that primarily drives the spread of infection in such epidemics. Our analysis suggests minimizing the movement of superspreaders during the lockdowns can be critical to the success of such lockdowns as it can delay/circumvent the onset of the switch.

Results

Country specific doubling rate of the infection is dependent on testing rate

To compare the dynamics of infection spread in different countries we first took the daily confirmed infection time course data from WHO and clustered the region-wise data according to their dynamic pattern. A hierarchical clustering algorithm (from the pheatmap package in R) is used to analyze the dynamics of around 100 countries selected (countries with at least 1000 infections per day in their

maximum infection spread phase). Figure 1A shows provinces in China and S. Korea, for instance, is clustered together as infection spread happened there at the earliest times (Figure 1A). Subsequently, infection in different parts of Europe/US is followed by the infection in other asian and south american countries (Figure 1A, color bar represents daily confirmed cases normalized to the maximum for each country ; see Figure S1A for an enlarged version representing certain regions in the heatmap). Next, we calculated the doubling rate from the time series of the countries. The doubling rate is defined as the inverse of the doubling time, i.e., how much time does the population take to double the number of infections (Figure 1B). Both figure 1A and figure 1B show how infection peaks and doubling rates vary across countries. Thus, for a better comparison of time evolution of the infection trajectories we superimposed them. In order to superimpose the trajectories, the doubling rate time courses for the different countries were aligned at the maximum doubling rate (SI for details) and linear correlations were calculated(Methods) between different timepoints of the aligned trajectories and numbers of tests/10000 population across these countries (Figure 1C). We observed that the normalized daily cases shows a positive correlation with the test rate around a time point of 16 days (Figure 1C, S1C, where 0 corresponds to time of maximum doubling rate), suggesting that the countries with higher test rates tend to have higher confirmed cases. On the other hand, the doubling rate after around 26 days displays a negative correlation with test rate (Figure 1D, S1B lower panel), suggesting that the countries with higher testing rate are able to reduce the rate of infection spread faster. However, during the start of the infection doubling rate is usually high for higher testing rate indicated by a positive correlation. Thus, higher testing rate (Figure S1B top panel) during the onset of infection allows more precise measurement of the infection spread which can be critical in containment of the infection through preventive measures such as quarantine. Notably the testing strategies in most of the countries are implemented under lockdown conditions. In order to project required enhancement in testing w.r.t to different extent of unlocks we next took recourse to a quantitative epidemic model and subsequently an agent based model.

A compartmental epidemic model is calibrated to obtain country specific infection parameters

From the group of 100 countries, we selected 50 representative countries, comprising a combination of infection stages and infected population size, and fitted their confirmed (Co), recovered(Re) and dead(De) population trajectories (Methods for details). The model contained five compartments namely susceptible (S), exposed (E), Infected (I), Quarantined (Q) and removed (R), where R contains both recovered and dead population²³. Figure 2A shows the structure of the implemented SEIQR model and

differential equations capturing the infection dynamics are explained in details in the methods section. The infection flow in the model starts when a susceptible person is exposed to the infection through transmission from an already infected person (rate parameter β). After exposure, the symptoms are exhibited within an incubation time; the exposed individual is infected (rate parameter α_1) and gets quarantined (rate parameter α_2); the quarantined individual either recovers (rate parameter γ_r) or dies (rate parameter γ_d) after a time delay, represented by a recovery or death rate. The confirmed fraction of the population by definition comprises [quarantined + recovered + dead] population and the infection (I) compartment represents the unidentified/unknown infection cases present in the population. An infected person once identified via testing is typically quarantined, hence in the model only the I but not Q further infects the S.

Dynamics of the system was captured by a set of 6 coupled ordinary differential equations (details in methods). The calibration data for each country comprises the time courses of the number of Co/Re/De and through model fitting we estimated the parameters that best explains the Co/Re/De trajectories simultaneously for a given country. Model fitting also includes a function that captures lockdown²⁴ where the lockdown is introduced in the model through a reduction in the transmission rate that follows an inverse sigmoid function (details in methods). The process of lockdown is controlled by three parameters - time of lockdown (start time of lockdown implementation), strength of lockdown (the extent of lockdown in a country, 0.1 would mean 90% lockdown) and the effectiveness of lockdown (how fast the maximum lockdown is achieved from the lockdown starting time point, determined by the steepness of the sigmoid curve).

Figure 2B shows the confirmed cumulative and daily cases, as well as the estimated lockdown extent and dynamics for four countries with different populations and stages of infection. Model fits for the 50 countries are shown in supplementary Figure S2. The average incubation time (average of $1/\alpha_1$) estimated in the fits is approximately 8.1 days which is close to the observed value²⁵. Our simplified model structure quantitatively captured the infection dynamics of countries with different approaches towards lockdown and subsequent unlock implementation. For instance, Australia (Figure 2A, top panel), after containing the first wave for a long duration, faced the second wave during unlock 1, were prompt to implement the subsequent lockdown and second wave of infection is contained too. The fitted lockdown function (Figure 2B, 3rd column) shows how the lockdown, unlock and lockdown values that best fits the data obtained for Australia where 0 and 1 in the y axis corresponds to complete and no lockdown, respectively. Rest three countries also have distinct dynamics and lockdown-unlock characteristics; for instance a continuously and rapidly growing number of confirmed cases in the USA

stems from an ineffective lockdown and subsequent unlocks. It can be noted that even in countries like Australia or Austria where lockdowns were relatively successful, plausibly > 90% effectiveness in implementation was achieved (Figure 2B, 3rd column, best-fit estimates of the lockdown efficiency) and rest of the population unaffected by the lockdown perhaps continues to spread the infection during the lockdowns. A recent landmark study indeed shows that a class of infected subjects called ‘superspreaders’ can account for a dramatic increase in the number of infected cases where a small fraction of superspreaders are responsible for the majority of infection reported in two states of India²². This suggests if a small fraction of the population spreads infections as a superspreader the overall effect of lockdown could be dramatically reduced. In a later section we quantitatively show how a small change in superspreader fraction can dramatically change the infection number.

Next, we obtain the dynamics of R_0 for different countries by utilizing the dynamics of the lockdown function. Our calculation shows a wide variety of responses among countries with respect to the effect of lockdown. Some countries are able to reduce the R_0 drastically after lockdown whereas in some countries impact of lockdown on R_0 remains relatively low (Figure 3A). Also, the test rate has a negative correlation with R_0 just after lockdown (Figure 3B), but this correlation is not significant. R_0 is presumably dependent on many other factors including demographics, medical facility and distribution of virus strains in the population, which can have distinct influences on the impact of implementation of testing in a given country/region. In order to systematically compare the influence of such factors on R_0 we collected several publicly available datasets of demographics, medical facilities and genome sequences in a country specific manner¹ (details in methods). R_0 values have significant correlation with some factors including median age, life expectancy, population size, doctors/10000 (Figure 3C, S3A). We also observe statistically significant correlation of R_0 with frequency of the clades in the population, specifically the L and GR strains shows significant influence on R_0 (Figure 3D, S3B). We next linearly adjusted the R_0 values (methods) using the most significant factors derived from the correlation analysis, namely median age, population size, doctors/10000 and frequency of the L clade which shows adjusted R_0 values exhibit much higher correlation with the test rate (Figure 3E). Here, we choose countries only with at least 10 complete genome sequences till the end of June 2020 for the correlation analysis which selects 39 countries out of the 50 countries. Next we specifically selected 5 representative countries with different transmission rates: high transmission rates (USA, India), low transmission rates (Australia, Austria) and high transmission rate and high quarantine rate (France) and studied the nature of infection spread in these different scenarios (Figure 3F shows countries with distinct transmission rates and quarantine rates).

Optimum changes in country-specific quarantine rates can effectively reduce the infection post lockdown

Typically, when lock down is removed, the transmission rate (β) is expected to increase. For our SEIQR model the R_0 value is given by $R_0 = \frac{\beta \rho}{\alpha_2 \rho_{max}}$ (see SI for derivation) where ρ captures the extent of lockdown ($\rho=1$ and $\rho = \rho_{max}$ represent full lockdown and full unlock respectively). From this expression we can calculate the extent by which the quarantine rate (α_2) must be increased in order to compensate for the effect of partial or full unlock. The fold change in α_2 to keep the R_0 value close to one is given by (SI for derivation)

$$F_\alpha = \frac{\beta \rho}{\alpha_2 \rho_{max}} \quad (1)$$

Here, $\rho = 1$ corresponds to fully lockdown case and gradual increase in value of ρ depicts partial unlock finally moving towards full unlock at $\rho = \rho_{max}$. The equation shows that the required fold change to compensate for the effect of lockdown removal would reduce as the extent of unlock reduces(Figure 4A). In fact, model fitting shows a higher transmission rate for the USA compared to Austria, Australia which asks for a higher quarantine rate for the same extent of unlock (Figure 4A). However, for instance for France, in spite of very high value of β the required increase in quarantine is low due to it's already high quarantine rate (Figure 4A). Better control on infection spread during lockdown removal would require increased quarantine rate which can only be achieved by rigorous contact tracing and testing. This particular scenario can be studied as a cost-benefit tradeoff where in one hand removal of lockdown benefits economic activities but at the same time more testing needs to be conducted to circumvent the enhanced rate of infection spread. If we assume that the maximum quarantine capacity of fold change in α_2 is F_c and investment cost in testing is λ per testing, the optimal fold change in α_2 would be (see SI for details)

$$F_\alpha^{opt} = F_c \frac{\left(\lambda \frac{F_c^2}{F_0}\right)^{\frac{3}{2}} + 1}{\left(\lambda \frac{F_c^2}{F_0}\right)^{\frac{1}{2}} + 1} \quad (2)$$

Figure 4B shows an illustration of the optimality in quarantine rate given a maximum capacity of quarantine rate. Finally from equation (1) the optimal lockdown removal (ρ_{opt}) is given by

$$\frac{\beta \rho_{opt}}{\alpha_2 \rho_{max}} = F_c \frac{\left(\lambda \frac{F_c^2}{F_0}\right)^{\frac{3}{2}} + 1}{\left(\lambda \frac{F_c^2}{F_0}\right)^{\frac{1}{2}} + 1} \quad (3)$$

Hence as the quarantine capacity F_c is low, the ρ_{opt} would be high (Figure 4C) which illustrates: as the quarantine capacity is low, the extent of lockdown should remain optimally high. Although higher F_c value is possible, total benefit would be low given the cost of testing. For instance, as the USA has a higher value of β , given a capacity of quarantine the optimally beneficial unlock would be low for the USA (Figure 4C) compared to Austria or Australia with much lower values of β compared to the USA and can afford a greater degree of unlock. Similarly, it also shows that the optimal extent of lockdown would increase if the cost of testing increases. Indeed, the extent of lockdown can be effectively manipulated by implementing a periodic lockdown-unlock cycle with a time interval of full unlock (T_1 followed by a time interval of full lockdown (T_2)) [20] within a time cycle of period $T = T_1 + T_2$. In this case one can show that in the allowed unlock period the average R_0 would be given by (SI for details)

$$\langle R_0 \rangle = \frac{\alpha_1 \beta}{\alpha_2 (\alpha_1 + \alpha_3)} \frac{1}{\rho_{max}} \left[(\rho_{max} - 1) \frac{T_1}{T} + 1 \right] \quad (4)$$

(here α_3 corresponds to the quarantine rate through testing in addition to quarantine rate α_2 of only symptomatic individuals; SI for details)

Equation (4) shows that the reproduction number during lockdown $\left(\frac{\alpha_1 \beta}{\alpha_2 (\alpha_1 + \alpha_3)} \frac{1}{\rho_{max}} \right)$ must be less than one to maintain the average R_0 less than one.

Frequency of the unlock can be connected to the extent of unlock through the equation below

$$\frac{T_1}{T} = \frac{\left(\frac{\rho}{\rho_{max}} - \frac{1}{\rho_{max}} \right)}{\left(1 - \frac{1}{\rho_{max}} \right)} \quad (5)$$

This quantifies the frequency of unlock required within a time N to implement a particular extent of overall unlock $\frac{\rho}{\rho_{max}}$ (Figure 4D). The corresponding unlock period to keep the average value at 1 is given by

$$\frac{T_1}{T} = \frac{\frac{\alpha_2 (\alpha_1 + \alpha_3)}{\alpha_1 \beta} \frac{1}{\rho_{max}}}{1 - \frac{1}{\rho_{max}}} \quad (6)$$

which represents allowed number of unlock days over a period time T to keep the average R_0 at equal to 1. Hence, as the effect of lockdown ρ_{max} is high the allowed unlock interval can also be longer. The equation further reveals that the allowed period of working days (T_1) can further be stretched by increasing the quarantine rate (α_3) (Figure 4E) involved with testing and quarantining the exposed subjects whereas small value of α_3 corresponds to the case when the testing rate is low assumed in fitting the data.

Agent based contact tracing model quantitatively connects testing rate and unlock percentage

As the ODE modeling framework used to study the dynamics of infection and quantitative calibration cannot model contact tracing and capture the number of testing required to achieve a certain quarantine rate α_3 , we have next deployed an agent-based model (ABM) to introduce testing through contact tracing such that a quantitative understanding on the required number of testing per unit population can be also be derived (details in methods) in a country specific manner. The agent-based stochastic version of the SEIQR model was simulated using country specific infection parameters obtained via calibrating the ODE model (methods). In the ABM model the infection spreads outward, starting with one infected agent at the centre, as shown in Figure S5. The selected model for the parameter values of the USA, for instance, shows that as the testing rate increases the maximum confirmed cases also increases (Figure 4F, qualitatively in agreement with data shown in figure 1C), but this also leads to containment of infection much earlier leading to a reduction in the reproduction number during different partial unlocks (Figure 5A). The results are comparable for other countries as well (Figure S4A-D). We next used an interpolation method based on cubic spline fitting (smooth.spline from R) to determine the minimum test rate required to keep the $R_0 < 1$ during each partial unlock as indicated in the Figure 5A. This calculation results in an monotonic increase in the required testing rate as the extent of unlock gradually progresses from full lockdown towards full unlock (Figure 5B). The exact values of required testing rates vary among countries for different partial unlocks depending on country specific transmission and quarantine rates (Figure 5B). It is noticeable that due to low transmission rate before lockdown, countries like Austria, Australia are estimated to be able to unlock by 80% with their current testing rates, whereas for the USA, the current testing rates need to be ramped up at least by 3 times to achieve an unlock of 80%. A 80% unlock corresponds to a periodic cycle of 10 working days followed by 5 days of lockdown (see Figure 4D). This suggests the USA with an increase its testing rate by at least 3 times can proceed to a stretch of 10 working days within a half month cycle but Austria/Australia could open their economy for 10 working days with their current testing rates. **In fact, this apparently projected connection between the agent based model and the analytical calculation is justified by the observation that the quarantine rate required to maintain R_0 at a value of 1 from the analytical calculation follows a linear scaling relation with the corresponding testing rate required to maintain R_0 at 1 from the agent based model (SI for details).**

A small fraction of superspreaders can dramatically increase infection spread during lockdown

Through the analysis presented above, we demonstrated that the burden on the testing rate capacity can be essentially reduced by keeping the transmission rate low (e.g. Austria) implemented through strict

standard measures of social distancing, wearing face masks etc. However, even during full lockdown, a small fraction of superspreading individuals²² can substantially dilute the diligent efforts exercised by majority of the population obeying strict lockdown rules. The superspreaders are defined as having low quarantine rate and high transmission rate compared to the regular spreaders. A recent detailed contact tracing analysis focusing on a couple of states in India demonstrates that during full lock down period a small fraction of superspreaders was responsible for transmitting 80% of reported infection, so even with rigorously planned full lockdown $R_0 < 1$ could not be achieved²². Such contribution of superspreaders may generally explain why in most countries full lockdowns could not contain the infection effectively. Using our agent-based contact-tracing model we simulated different scenarios where a population comprises different proportions of superspreaders.

Our simulations suggest that a smaller increase in the fraction of superspreaders can drastically increase the number of confirmed cases (Figure 5C & Figure S4E, simulated with parameter sets specific to India) and also the R_0 (Figure 5C-inset) when the testing rate is kept high to best track the spread. A small number of superspreaders (5%) contribute to the spreading of the majority of the infection (~70% when testing rate in 9/10000, Figure 5D). It can be noted that a fraction of total infection for 5% superspreader fraction is marginally less compared to the condition when superspreader composition is 10% in the population (Figure 5D). This indicates a small number of superspreaders above a critical fraction is sufficient to turn on a switch-like infection spread beyond which an increase in the number of superspreader fraction in the infected population may not be as detrimental. The total number of confirmed cases for 5% or 10% superspreaders are also comparable and are distinct from the scenario when the superspreader fraction is 0% (Figure S4E, shown for India). In fact, we observed that the fraction of superspreaders required to have a 50% contribution in total infection over a period of time can be as low as 1% for India, whereas for Austria 3% superspreader is needed for the same effect (data not shown); here, Austria owing to a lower intrinsic transmission rate is able to tolerate more superspreaders whereas superspreader movement in high transmission rate countries like India (or USA) can result in dramatic increases in overall infection. In order to better understand the mechanism for this switch-like behavior, especially at high quarantine rate, we conducted deterministic simulations at different values of transmission rates (β). The total confirmed cases increases in an ultrasensitive switch-like manner at a high quarantine rate (α_2 , the mechanistic equivalent of testing rate in the agent based model) while the increase is more gradual at low α_2 (Figure 5E). This switch can be understood as a consequence of the inherent systems-level positive feedback loop that drives the infection spread in an epidemic outbreaks where testing and quarantining are implicit components. While rigorous

testing and quarantining is a necessary measure to contain the infection it can only delay the onset of the exponential phase of infection spread and movement of superspreaders helps reaching the exponential phase faster. Our agent-based model simulation show drastic increase in R_0 of total population with an increase in the fraction of superspreaders (Figure 5C). In the context of the ODE model, the superspreaders and regular spreaders would typically have β value above and below the switch threshold (Figure 5E), respectively, during the lockdown. This will lead to an overall dramatic increase in the number of infected populations owing to the superspreaders' contribution. Thus, in order to leverage the benefit of augmented testing capacity, tracing the movement of superspreaders appears as crucial as implementation of other standard social distancing measures.

Discussion

The COVID19 epidemic has disrupted normal life in an unprecedented manner in almost every corner of the world. Similar such outbreaks, with different infective capacity were reported earlier, for instance, the basic reproduction numbers or R_0 for COVID19 is comparable to the SARS-cov^{26,27} and much higher than the MERS infection^{28,29}. In this study we quantitatively explored the COVID19 outbreak and its relation with quarantine measures in several countries using both deterministic and agent-based models coupled to analysis of country specific metadata. We first employed a deterministic (ODE based) SEIQR epidemic model and fitted the trajectories from 50 countries at different stages and starting time(day) of infection. We then built an agent based stochastic model³⁰ to implement contact tracing by utilizing the best-fit parameters from the ODE based SEIQR model and implement them in the agent-based model while ensuring a quantitative scaling between both models. The primary goal of utilizing the two distinct modeling approaches is to connect the testing rate to the unlock process since contact tracing can not be implemented in ODE based epidemic models. Earlier studies discuss the limitations of ODE based models and agent based models³², but studies connecting both modeling approaches especially in the context of the COVID19 epidemics to understand the relation between testing rate and unlock measures are not reported in the literature. Here, feeding the country specific epidemic parameters from the ODE model into the agent based model we derive country specific optimal testing-rate through contact tracing. Such analysis aims to guide different degrees of unlock to open up economies in different countries since long term lock down is impractical and detrimental to the economy of any country³¹.

Typically, the actual infected number of people is expected to be higher than the sampled one's and limitations like that pose challenges for devising accurate mathematical models³². Actual susceptible population size is thus mostly unknown and it could often be different from the tested population size,

so the number of infected people from the data only captures the tested sample when the total number of tests are less than the total population of a region/country. This is also one of the reasons why increased testing rate is critical in better capturing the magnitude of the infection (and not only the nature of the dynamics, the dynamic features of the infection spread can be robustly captured even with a relatively smaller tested population size) that can lead to devising more accurate epidemic models with better predictive capacities in early stages of the infection. In addition, because of the asymptomatic nature of the infection³³, the need to quarantine the maximum number of infectious people through testing is equally important to successfully contain the infection spread. Apart from the testing rate, other factors like immunization³⁴, age³⁵, sex³⁶ medical facility and specific virus strains³⁷ may play important roles both in transmission and fatality but testing rate is the only variable amongst all these critical factors that can be manipulated immediately post infection via human intervention.

In countries like India where population density and size outmatch the healthcare infrastructure, compared to Austria for instance where population size and health infrastructure are better matched, an early implementation of lockdown was critical in slowing down the spread of the infection. Similarly, in the early days of COVID19 spread, delay in lockdown implementation had a catastrophic impact in Italy or Spain despite adequate health infrastructure per capita, highlighting the importance of implementation of early and rapid measures post infection irrespective of healthcare infrastructure. This is further highlighted where unlock measures chosen by countries like India and USA immediately witnessed dramatic increase in infection spread.

Reopening the economy is also an impending necessity in all countries under lock down. Here, utilizing quantitative knowledge on infection and lockdown dynamics from the ODE model and feeding the country-specific best-fit parameters into the agent-based model we propose strategies for optimal testing and quarantining of the infected subjects to maintain $R_0 < 1$ during the unlock periods. We quantitatively demonstrate the quarantine and testing that can be optimally adjusted (post lockdown) to minimize the infection spread. We demonstrate that the effective increase in testing mediated quarantine measures has to be country specific, primarily depending on the transmission and quarantine rates of a country. Already, many countries have taken initiatives to accentuate the testing capacities either through developing more testing facilities¹⁴⁻¹⁶ or through designing efficient pool testing algorithms^{38,39}. Our study can additionally help facilitate more informed planning of such testing processes where, for instance, if a goal of 30% unlock of economy is set by a given country(or a province in a country), then the respective increase in testing rates to achieve such unlock goals can be predefined as the minimal testing rate target. In an unprecedented efficiency multiple vaccines are

rolled out now, which will eventually aim to contain the COVID19. But it is still relatively unclear how the rapidly evolving variants of the virus will impact the efficacy of different types of vaccines as the vaccines have distinct mechanisms of target engagement. Further, many developing or under developed countries may not have access to such wide spread immunization programs to vaccinate their entire population till at least 2022⁴⁰. Thus systematic planning of testing and designing of lockdown and unlock measures remain as a key factor in containing the current spread of COVID19 and in specific and perhaps in case of similar such epidemics in the future.

Periodic lockdown²¹ is another efficient way to implement partial lockdown which best works if the R_0 value during the lockdown is much lower than 1 maintained typically via rigorous testing and quarantining during the lockdowns. However in many developing as well as developed countries, it became almost impossible to limit the mobility of people during lockdown^{11,25} thus maintaining a low R_0 during lockdown remains a challenge. A small unidentified fraction of the exposed population during the unlocks can potentially remain unidentified due to the long incubation period characteristic to COVID19. Indeed, we saw a higher surge of a second wave all over Europe that qualitatively resembles the second wave of Spanish flu pandemic of 1918⁴¹. A recent landmark study highlights how a small fraction of unidentified superspreaders dramatically accelerated the spread of infection in two states of India even during nation-wide lockdown²² and diluted India's advantage of early lockdown. Looking at the trend of infection dynamics across the world this could be true to multiple other countries as infection numbers are growing constantly even during the lockdowns. [Here we quantitatively compare for the first time the significance of relative number of superspreaders in a country specific manner; for instance our agent based simulations during the lockdown suggest 1% superspreader can contribute to a 50% increase in total infected number in a given timescale in countries like India whereas in countries like Austria 3% of the infected population has to be superspreader to obtain the same effect.](#) Thus the discovery of superspreader mediated infection spread via contact tracing can be a critical component of lockdown implementation itself which further underscores the advantages of our proposed strategy of optimally augmenting the test rate through contact tracing. Additionally, the recent [emergence of a few more new strains may also have accelerated the upsurge of confirmed cases⁴⁵ in addition to contribution from unlocking and reopening the economies and the superspreaders contribution during the lockdown .](#)

In this study, using an approach of cost-benefit optimization we derived a data-driven strategy that aims to achieve $R_0 < 1$ during unlocks which will perhaps also reduce the severity of uncontrolled future outbreaks. A general property of epidemic outbreaks such as COVID19 is that they are fueled by the

strength of an implicit positive feedback loop connecting the infected population to the susceptible population. Such positive feedback can promote switch-like responses especially when the testing/quarantine rates are high. The switch-like response ensures that for a high testing/quarantine rate infection spread can be kept below the switch threshold even for significantly higher values of infection transmission rate but for low quarantine rates exponential rise in infection can occur for a much lower transmission rate. Our simulations suggest a small fraction of superspreaders (with high β) can exploit the switch-like response capacity of the system to trigger a major increase in infection number during the lockdowns. As the superspreader fraction increases beyond a threshold even for a high testing rate scenario a sudden increase in the number of confirmed can result due to overall change in strength of the positive feedback loop. Thus, the containment of superspreaders during the lockdown would be critical to maintain a low R_0 such that subsequent unlocks can be implemented in an optimal manner, especially in the densely populated cities where the transmission probability is higher. In fact, recent upsurge of infection spread in several regions even after launching of vaccination programs⁸ reiterates the warning that our complacency in obeying all social distancing and quarantine rules out of desperation to return to pre-pandemic life may incur fatal consequences. Finally, if this COVID19 pandemic is successfully contained in the coming months due to improvements in the efficacy of vaccines, the strategy we propose here may still facilitate better containment of future such outbreaks.

Materials and Methods

Calculation of doubling rate

The doubling rate is calculated at each time point over the trajectory by taking the daily cases at that time and next day. For an exponential growth of the infection, $N(t) = N(t + 1)exp(r_d)$ where $N(t)$ depicts the number of infected people at time t

$$r_d = \log\left(\frac{N(t+1)}{N(t)}\right)$$

Here $t=1$ day. The values in the heatmap in Figure 1B are displayed in an exponential scale of doubling rate to avoid negative values. So a doubling rate of 1.4 in exponential scale means the doubling time of $\frac{1}{\log(1.4)} = 3$ meaning it takes around 3 days to double the number of infected people.

Correlation analysis

For the correlation and significance of the correlation we used `cor.test()` function in R. The function uses a t-distribution statistics to calculate the p-value of the correlation. The t value of a pair of random

numbers with correlation r and number of points n is give by $t = \frac{r\sqrt{n}}{\sqrt{1-r^2}}$

The t-distribution provides the probability of t values for the null hypothesis that the mean correlation of the two random variables is zero. If the calculated correlation value falls on the tail of the t-distribution the, the correlation is significant. The exact p-value is calculated from the t-distribution table of n degrees of freedom. The calculations are all done by the `cor.test()` package in R.

The SEIQR model: The model comprises of susceptible(S), Exposed(E), Infected(I), Quarantined(Q), Removed(R which contains two compartments ‘recovered’ and ‘dead’)

The equations are

$$\begin{aligned}\frac{dS(t)}{dt} &= -\beta * \rho(t) * \omega(t) * S(t) * I(t) \\ \frac{dE(t)}{dt} &= \beta * \rho(t)\omega(t)S(t)I(t) - \alpha_1 E(t) \\ \frac{dI(t)}{dt} &= \alpha_1 E(t) - \alpha_2 I(t) \\ \frac{dQ(t)}{dt} &= \alpha_2 I(t) - \gamma_d Q(t) - \gamma_r Q(t) \\ \frac{dR(t)}{dt} &= \gamma_r Q(t) \\ \frac{dD(t)}{dt} &= \gamma_d Q(t)\end{aligned}$$

Where $S(t)$, $E(t)$, $I(t)$, $Q(t)$, $R(t)$ and $D(t)$ are the susceptible, exposed, infected, quarantine, recovered and dead population at time t respectively. $\rho(t)$ is the lockdown function and $\omega(t)$ is the time when infection starts in a given country after the first day of detection of the infection date as reported in WHO website (hence simulation start time corresponds to 22nd Jan, 2020 based on the WHO report).

The lockdown function $\rho(t)$ is given as

$$\rho(t) = \frac{1}{1 + \left[\frac{K}{1 + \exp(tc - t)} \right]^S}$$

where, K =Effect_lockdown_Country, S = strength of lockdown and tc = time of lockdown

$\rho(t)$ varies between 0 and 1(no lockdown) where 0 (full lockdown). $\omega(t)$ ensures that the model for a specific country is switched on when the infection begins in that country, hence if the detected case in a country is 40 days after Jan 22nd, the model for that country is switched on 35 days post Jan 22nd

(assuming mean incubation time of 5 days), during the calibration.

The lockdown is opened by modifying the $\rho(t)$ function such that $\rho(t)$ returns to 1 from its lockdown status to no-lockdown(1) status in a designated time .

Model calibration

Model calibration involves minimizing an objective function that gives best fit parameter sets for confirmed, recovered and dead populations for a given country, simultaneously. We fitted the time series provided by JHU CSSE at github⁴³ to the SEIQR model developed in the study and minimized the objective function using the [nonlinear least square fitting algorithm \(lsqnonlin\) from MATLAB](#).

It returns $x = \text{lsqnonlin}(\text{fun}, x_0, \text{lb}, \text{ub})$, where x is a vector/matrix of variables whose values are to be determined, for instance, in our case x comprises the country specific infection parameters to be estimated via the fitting the model to the observed infection dynamics for a given country of interest. The function 'fun' contains the residuals or sum of squares of the difference between model value and data point for a given time point; the residuals are evaluated for each time point for which data point is available. x_0 is the initial parameter vector that comprises the initial guesses of the model variables and ub and lb respectively represent upper bound and lower bound vectors corresponding to each value in x_0 . For instance, to calibrate the confirmed population trajectory for the USA the fitting algorithm minimizes the sum of residuals for $[\text{Quarantined_USA} + \text{Recovered_USA} + \text{Dead_USA}]$ between both model and data. This exercise was carried out for all the 50 countries individually to obtain the country specific infection parameters.

Analysis of the virus genome sequences for clade distribution

We first collected the individual sequenced genome from individual countries and their corresponding clades as reported⁴³. Each genome in a particular country belongs to one of the seven clades (G,GH,GR,L,S,V.O) as described⁴³. From the data, we determine the frequency of each clade in individual countries by finding the number of sequences belonging to that clade divided by the total number of sequences. To calculate the correlation, out of the analysed 50 countries, we only considered countries having at least 10 sequences in the database.

Adjusted R_0 values

The R_0 values obtained by fitting are adjusted with respect to the significant parameters from the correlation analysis described in Figure 3C,D and S3A,B. The adjustment factor is characterised as

$$A = \frac{1}{\text{populationsize}} * \text{median} \frac{\frac{\text{age*doctors}}{1000} * \text{nurses}}{\frac{1000}{\text{frequencyofGRclade}}} * 1$$

The factors which are positively correlated with R_0 (population size, frequency of GR clade) are placed in the denominator and factors having negative correlation values are placed in the numerator. The adjusted R_0 values are quantified as

$$R_0^{\text{adjusted}} = \frac{R_0}{A_n} \text{ where } A_n = \frac{A}{\text{max}}(A)$$

Agent based stochastic SEIQR model

We utilized the model structure and calibrated parameters in the deterministic SEIQR model, as described in the previous section to build an agent based stochastic version of the same³⁰. In this model, the individual agents are assumed to be located on a two dimensional lattice of dimension 300×300 . The points on the lattice represent susceptible agents in the population and each individual on the lattice is surrounded by four nearest neighbors at a minimum distance of one. Here, we considered a radius of seven as the closest contacts of every agent on the lattice. Thus, every individual agent would be in contact with a maximum of 28 neighbors. The initial patient at zero time is located exactly at the center having a location of (150,150) on this two dimensional lattice. Infection spreads through persistent close contacts with the nearest neighbors. At every time step, all the exposed and the infected agents are selected and 28 neighbors of the individual get exposed to the infection with a probability β . The value β_{ODE} in the ODE model is defined as the number of individuals a particular person infects per unit time. In the stochastic model each individual would infect $28 \times \beta$ people on an average per unit time. Thus, the conversion between transmission rates of the ODE and stochastic model is given by $\beta = \frac{\beta_{ODE}}{28}$. Here we assumed that both the asymptomatic/exposed and symptomatic patients are capable of spreading the infection. The exposed patient in turn exhibits symptom with a probability α_1 and the symptomatic infected patient would be further quarantined with a probability α_2 . The quarantined patient finally either recovers or dies with a rate γ . The susceptible, exposed, infected, quarantined and recovered agents on the lattice are assigned values 0,1,2,3 and 4 respectively which allow us to track the dynamics of each component separately. The parameters values were taken from the fitted values of the ODE based model for a particular country to simulate the real scenario in that country. In order to introduce diagnostic testing, we identified all the symptomatic as well as the quarantined individuals from the pool at a particular time point and selected a fraction (F_{test}) of the identified patients for contact tracing. The contacts of each selected individual were traced and diagnostic tests were conducted on them. The close contacts showing positive results

were finally quarantined. This procedure is repeated for all the selected individuals. The status of the whole population is updated and the next time step is continued. The simulation was performed for 200 time points. In order to increase the number of testing, the value of F_{test} were increased. However, we assumed that the testing procedure only commences after at least 100 individuals are already infected in the population in order to introduce an initial delay in responding to the situation. For every infected person, we also keep track of the number of individuals who are getting infected from that particular infected person until that person is quarantined. The number of people infected by the time defines the reproduction number for that individual. The R_0 value for the whole population is quantified by taking average over the reproduction numbers for all the persons in the population.

Author contribution: US and BG conceived the study. US built the epidemic model, developed fitting and prediction strategies and optimized the parameters. BG developed the agent based model, performed the stochastic simulations, analytical calculations, genome sequence analysis and regression analysis. US and BG analyzed the results and wrote the manuscript.

Conflict of interest: The authors declare no conflict of interest.

Acknowledgement

Authors thank Department of Biotechnology (No. BT/RLF/Re-entry/32/2017), Government of India for funding.

Conflict of interest: The authors declare no conflict of interest.

References

1. World Health Organization, *Coronavirus Disease 2019 (COVID19) Situation Report – 66* (WHO,2020);https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200326-sitrep-66-COVID19.pdf?sfvrsn=9e5b8b48_2
2. R. Li, C. Rivers, Q. Tan, M. B. Murray, E. Toner, M. Lipsitch, The demand for inpatient and ICU beds for COVID19 in the US: lessons from Chinese cities. *medRxiv* 2020.03.09.20033241 [Preprint].
3. “Not a wave, a tsunami.’ Italy hospitals at virus limit” AP NEWS, 13 March 2020
4. “COVID19 infections rise in New York with peak weeks away.” AP NEWS, 25 March

2020;<https://apnews.com/7c7563cb82626a4042797c6aa6da260a>.R. M. Anderson.

5. DRAFT landscape of COVID19 candidate vaccines – 20 April 2020;
<https://www.who.int/blueprint/priority-diseases/key-action/novel-coronavirus-landscape-ncov.pdf>
6. T. D. Hollingsworth How will country-based mitigation measures influence the course of the COVID19 epidemic? (2020).doi:10.1016/S0140-6736(20)30567-5pmid:32164834
7. Q. Bi, Y. Wu, S. Mei, C. Ye, X. Zou, Z. Zhang, X. Liu, L. Wei, S. A. Truelove, T. Zhang, W. Gao, C. Cheng, X. Tang, X. Wu, Y. Wu, B. Sun, S. Huang, Y. Sun, J. Zhang, T. Ma, J. Lessler, T. Feng, *Epidemiology and Transmission of COVID19 in Shenzhen China: Analysis of 391 cases and 1,286 of their close contacts. medRxiv* 2020.03.03.20028423 [Preprint].
doi:10.1101/2020.03.03.2002842
8. Moore S, Hill EM, Tildesley MJ, Dyson L, Keeling MJ. 2021 Vaccination and non-pharmaceutical interventions for COVID-19: a mathematical modelling study. (2021) *Lancet Infect Dis.* 18:S1473-3099(21)00143-2.
9. Bartik, Alexander and Bertrand, Marianne and Cullen, Zoe and Glaeser, Edward L. and Luca, Michael and Stanton and Christopher (2020) The Impact of COVID-19 on Small Business Outcomes and Expectations *Proc. Natl. Acad. Sci. USA* 117 (30) 17656-17666
10. Patrick G. T. Walker (2020) The impact of COVID-19 and strategies for mitigation and suppression in low- and middle-income countries. *Science* 369 413-422

11. Joakim A. Weilla,1 , Matthieu Stiglerb , Olivier Deschenesc and Michael R. Springbornd (2020) Social distancing responses to COVID-19 emergency declarations strongly differentiated by income *Proc. Natl. Acad. Sci. USA* 117 (33) 19658-19660;
12. Giovanni Bonaccorsia et al. (2020) Economic and social consequences of human mobility restrictions under COVID-19. *Proc. Natl. Acad. Sci. USA* 117 (27) 15530-15535
13. Max Roser, Hannah Ritchie, Esteban Ortiz-Ospina and Joe Hasell (2020) - "Coronavirus Pandemic (COVID-19)". *Published online at OurWorldInData.org*. Retrieved from: '<https://ourworldindata.org/coronavirus>' [Online Resource]
14. Tromberg BJ, et al. (2020) Rapid scaling up of COVID-19 diagnostic testing in the United-States: the NIH's RADx Initiative. *N Engl J Med* 383:1071-1077
15. <https://www.who.int/india/news/feature-stories/detail/how-india-scaled-up-its-laboratory-testing-capacity-for-covid19>
16. <https://news.cgtn.com/news/2020-06-24/China-s-daily-COVID-19-testing-capacity-hits-3-78-million-RAa1yHsBYk/index.html>
17. Robert Marsland III and Pankaj Mehta (2020) Data-driven modeling reveals a universal dynamic underlying the COVID19 pandemic under social distancing *arXiv:2004.10666*[q-bio.P].
18. Kissler SM, Tedijanto C, Goldstein E, Grad YH and Lipsitch M (2020) Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period *Science*, 14 eabb5793.
19. Yubei Huang , Lei Yang , Hongji Daia , Fei Tiana , and Kexin Chena (2020) Epidemic situation and forecasting of COVID19 in and outside China DOI:[10.2471/BLT.20.255158](https://doi.org/10.2471/BLT.20.255158).
20. [Rajesh Singh](#) and [R. Adhikari](#) (2020) Age-structured impact of social distancing on the COVID19 epidemic in India *arXiv:2003.12055*.
21. Omer Karin *et al.* (2020) Adaptive cyclic exit strategies from lockdown to suppress COVID-19 and allow economic activity. *medRxiv* 2020.04.04.20053579v1
22. Laxminarayan R. *et al.* (2020) Epidemiology and transmission dynamics of COVID-19 in two Indian states. *Science* 370 691-697
23. J. Jumpen W et. al. (2020) A SEIQR model for pandemic influenza and its parameter identification *International Journal of Pure and Applied Mathematics* 52(2).
24. Jonas Dehning et al. (2020) Inferring change points in the spread of COVID-19 reveals the effectiveness of interventions. *Science* 369(6500):eabb9789

25. Qun Li et al. (2020) Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia *N Engl J Med* 382:1199-1207
26. Riley S., Fraser C., Donnelly C. A., Ghani A. C., Abu-Raddad L. J., Hedley A. J., et al. Transmission dynamics of the etiological agent of SARS in Hong Kong: impact of public health interventions. *Science*. 2003; 300(5627): 1961-6
27. Lipsitch M., Cohen T., Cooper B., Robins J. M., Ma S., James L., et al. (2003) Transmission dynamics and control of severe acute respiratory syndrome. *Science* 300(5627): 1966-70,
28. Chowell G., Abdirizak F., Lee S., Lee J., Jung E., Nishiura H., et al.(20215) Transmission characteristics of MERS and SARS in the healthcare setting: a comparative study. *BMC medicine* 13: 210, DOI: 10.1186/s12916-015-0450-0.
29. Kucharski A. J., Althaus C. L. (2015) The role of superspreading in Middle East respiratory syndrome coronavirus (MERS-CoV) transmission. *Euro Surveill*. 20(25): 14-8,
30. Erik Cuevas (2020) An agent-based model to evaluate the COVID-19 transmission risks in facilities. *Computers in Biology and Medicine* 121 10382
31. <https://www.economist.com/leaders/2020/04/30/life-after-lockdowns>
32. Andrea L. Bertozzi, Elisa Franco, George Mohler, Martin B. Short, and Daniel Sledge (2020) The challenges of modeling and forecasting the spread of COVID-19. *Proc. Natl. Acad. Sci. USA* 2020 117 (29) 16732-16738;
33. Seyed M. Moghadas et al. (2020) The implications of silent transmission for the control of COVID-19 outbreaks *Proc. Natl. Acad. Sci. USA* 117 (30) 17513-17515
34. Luis E. Escobar, Alvaro Molina-Cruz, and Carolina Barillas-Mury (2020) BCG vaccine protection from severe coronavirus disease 2019 (COVID-19). *PNAS* 2020 117 (30) 17720-17726

35. John P.A. Ioannidis, Cathrine Axfors,b,c and Despina G. Contopoulos-Ioannidis (2020) Population-level COVID19 mortality risk for non-elderly individuals overall and for nonelderly individuals without underlying diseases in pandemic epicenters. *Environ Res.* 188: 109890.
36. Takahashi, T., Ellingson, M.K., Wong, P. *et al.* Sex differences in immune responses that underlie COVID-19 disease outcomes. *Nature* **588**, 315–320.
37. Korber et al. (2020) Tracking Changes in SARS-CoV-2 Spike: Evidence that D614G Increases Infectivity of the COVID19 Virus. *Cell* 182, 812–827
38. Sabyasachi Ghosh et al. (2020) Tapestry: A Single-Round Smart Pooling Technique for COVID-19 Testing. *medRxiv* doi: <https://doi.org/10.1101/2020.04.23.20077727>
39. Aragón-Caqueo D, Fernández-Salinas J, Laroze D. Optimization of group size in pool testing strategy for SARS-CoV-2: A simple mathematical model. *J Med Virol.* 2020 Oct;92(10):1988-1994.
40. Aisling Irwin. What it will take to vaccinate the world against COVID-19 (2021) *Nature* 592, 176-178 . doi: <https://doi.org/10.1038/d41586-021-00727-3>.
41. John M. Barry,¹ Cécile Viboud,² and Lone Simonsen (2008)Cross-Protection between Successive Waves of the 1918–1919 Influenza Pandemic: Epidemiological Evidence from US Army Camps and from Britain. *J Infect Dis.* 2008 Nov 15; 198(10): 1427–1434.
42. <https://www.hindustantimes.com/india-news/fake-news-led-to-migrants-exodus-mha/story-AEYej7C7jaLakuQ2dCIcML.html>
43. Dong E, Du H, Gardner L. An interactive web-based dashboard to track COVID-19 in real time. *Lancet Inf Dis.* 20(5):533-534.
44. Elbe, S., and Buckland-Merrett, G. (2017) Data, disease and diplomacy: GISAID’s innovative contribution to global health. *Global Challenges*, 1:33-46.
45. Weisblum Y, Schmidt F, Zhang F, et al. (2020) Escape from neutralizing antibodies by SARS-CoV-2 spike protein variants *eLife* 9:e61312.

Figure legends

Figure 1: The Covid-19 outbreak in different countries and its relation with the testing rate. (A) The heat map displays the clustered dynamics for around 100 different countries for the daily cases normalized to the maximum for each country (Days correspond to 22nd January, 2020 and to 15th August, 2020). The dendrogram is based on hierarchical clustering of the time traces (B) The heat map represents the doubling rate as a function of time clustered according to the dynamics of the time traces as in (A). The results are shown in exponential scale. (C) The Pearson correlation coefficient of test rate with the daily confirmed cases at different time points along the dynamics are shown. The time traces are aligned w.r.t the maximum doubling rates. 0 on the x-axis corresponds to the maximum doubling rate for all the countries. (D) The Pearson correlation coefficient of test rate with the doubling rate at different time points along the dynamics aligned as described in (C).

Figure 2: The description of the SEIQR model and calibration for different countries. (A) Schematic representation of the SEIQR model which contains susceptible, exposed, infected, quarantined, recovered and dead compartments. The lockdown is implemented through a sigmoid function as shown in the methods section. The quarantined, recovered and dead cases together comprise the confirmed cases. The arrow from infected to susceptible represents the positive feedback that fuels the infection spread in the population which is negatively regulated by the lockdown, as indicated (B) The SEIQR models fit the data for 4 representative countries as indicated for the cumulative confirmed and daily confirmed cases. The number of days in X axis corresponds to the time course data available in JHU CSSE [41] where 0 corresponds to 22nd January, 2020 and the end time point corresponds to 30th October, 2020. The lockdown function shows the extent of lockdown with time where 0 = full lockdown and 1 = no lockdown.

Figure 3: The reproduction number (R_0) displays correlation with the testing rate across countries. (A) The barplot shows R_0 for the fitted 50 countries after lockdown. (B) The scatter plot describes correlation between the test rate and R_0 after lockdown for the 39 countries with at least 10 complete genome sequences as of end of June described in the main text. (C) The statistical significances of the dependence of R_0 on different demographic and medical facilities factors are shown based on linear regression based on the same 39 countries (D) The same statistical significance of dependence of R_0 on the frequencies of the clades (G,GR,GH,L,O,S,V) are shown. (E) The scatter plot indicates the correlation between the adjusted R_0 values and the test rates. (F) The values of the

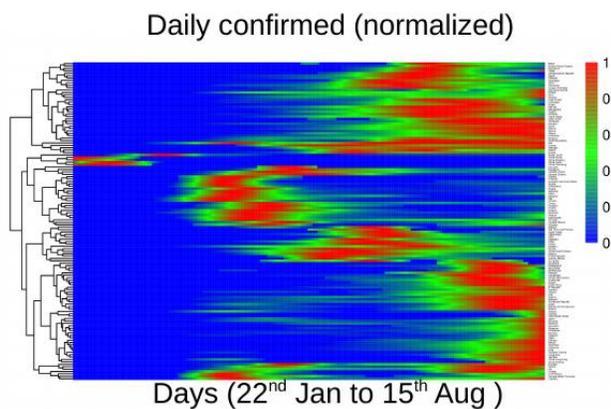
fitted transmission rates and quarantine rates for all the 50 countries are plotted and the representative countries for consideration of the agent based model simulation are shown in colors red, green and blue.

Figure 4: The optimization of trade-off between testing rate and extent of unlock. (A) The relationship between the extent of unlock and the corresponding quarantine rate required to keep R_0 values equal to one for five countries as indicated. The y-axis is plotted in log-scale. (B) The cost and benefit curves as a function of quarantine fold change for the countries. The cost benefit is optimized at a value F_α^{opt} for a maximum capacity F_c of the fold change in quarantine rate. (C) The optimum extent of unlock calculated from F_α^{opt} as a function of the maximum quarantine capacity F_c . Both the axes are plotted in log-scale. (D) The implementation of the extent of unlock as a function of the periodic unlock for the different countries are indicated. Number of working days represents the working day cycle in 15 days time. (E) The quarantine required to be increased by testing to keep the R_0 value below one and the corresponding the working days allowed. (F) The time traces of the confirmed daily cases of agent based simulation based on parameter values with full unlock for the USA at different daily tests/10000 as indicated.

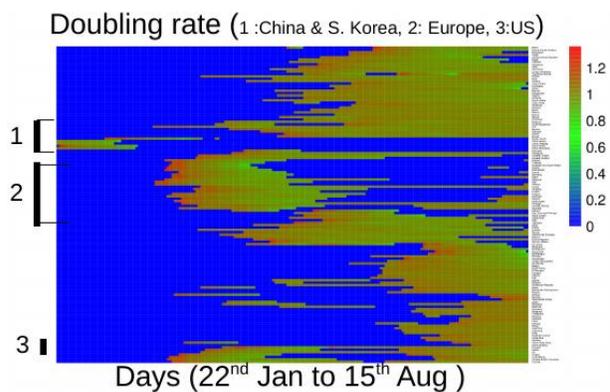
Figure 5: The testing rate and fraction of superspreaders in the population influence the infection rate simulated by the agent based model (A) The R_0 values based on the agent based model with parameter values of the USA as a function of test/10000 for different extents of unlock as indicated. The other four countries are shown in supplementary figure S4. The dashed horizontal line exhibits the $R_0 = 1$ line, intersection of the line with the curves quantifies the test/10000 required to maintain $R_0 = 1$. (B) The allowed extent of unlock and the corresponding test/10000 required to keep the R_0 value less than one calculated from the agent based simulation for the different counties with indicated current test rate. (C) The total confirmed cases as a function of percentage of superspreaders simulated for parameter values of India at different testing rates as indicated. (The inset shows the corresponding R_0 values at different percentages of superspreaders). (D) The barplot indicates the contribution of superspreaders and normal spreaders over the total number of confirmed cases as a function of the percentage of superspreaders at 3 different testing rates 0.2/10000, 2/10000 and 9/10000, as indicated. (E) The relationship between the transmission rate and the total confirmed cases at two different values of quarantine rate as indicated for two representative countries, USA and India.

Figure 1

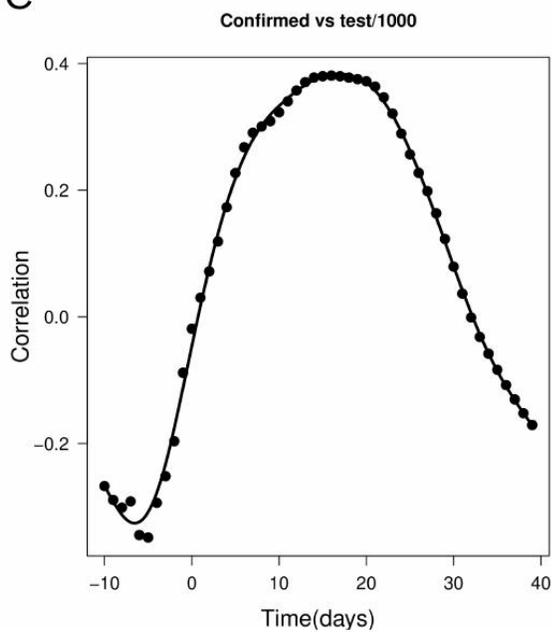
A



B



C



D

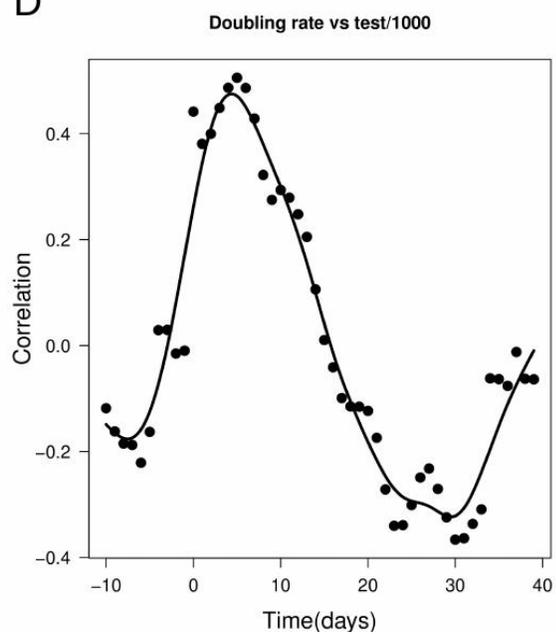


Figure 2

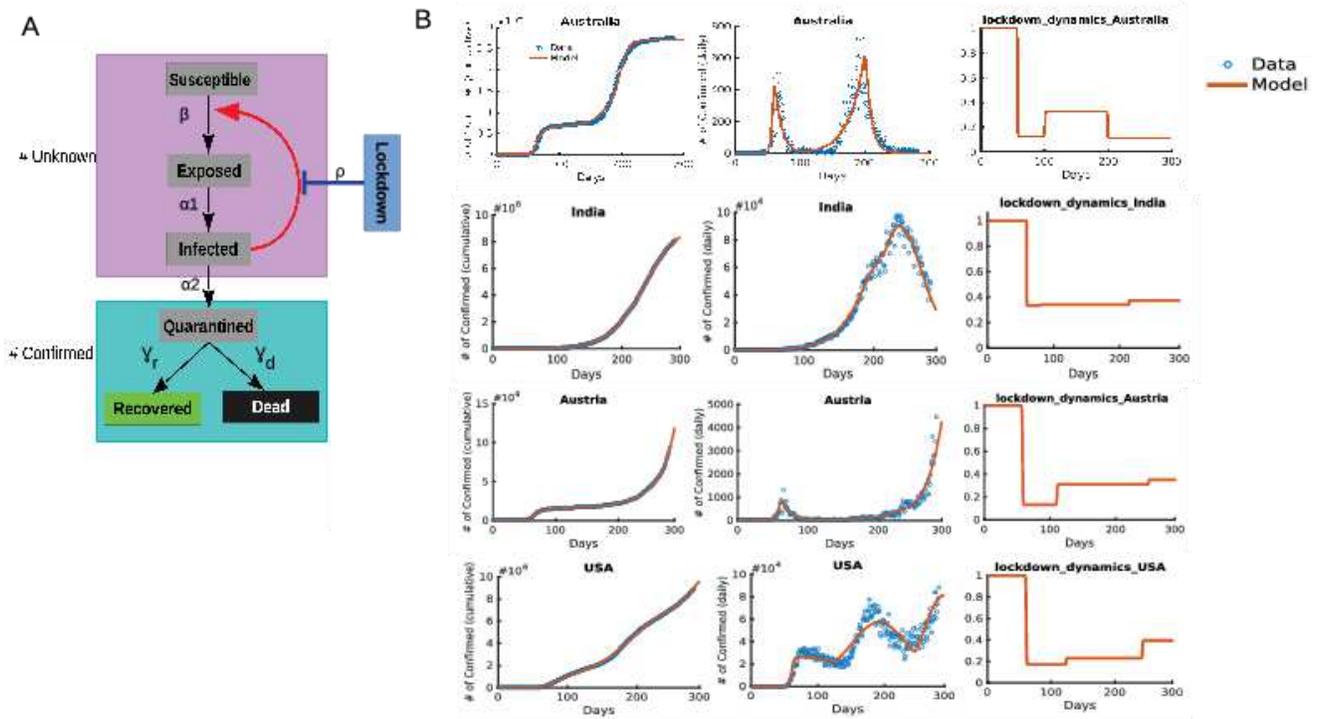


Figure3

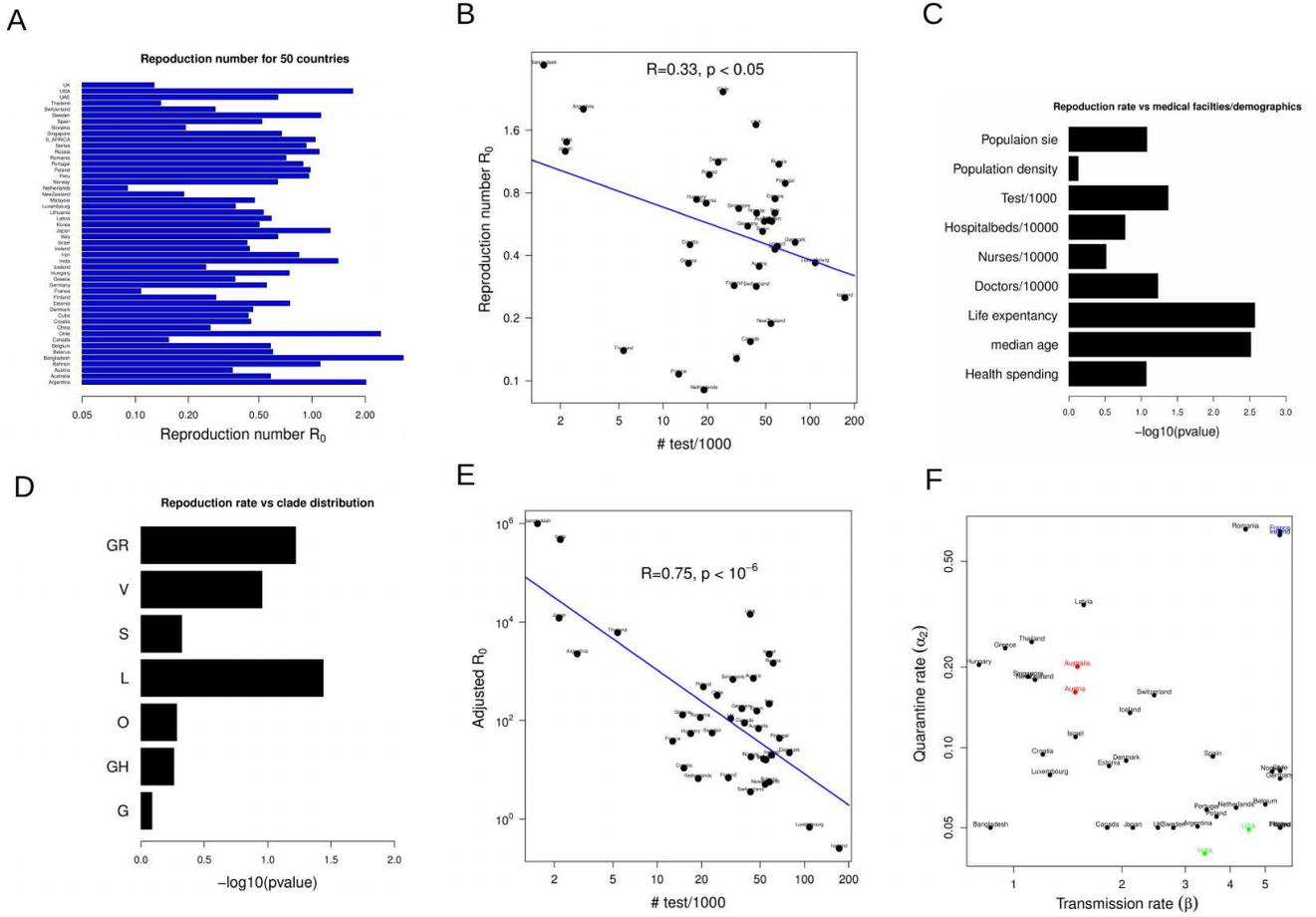


Figure 4

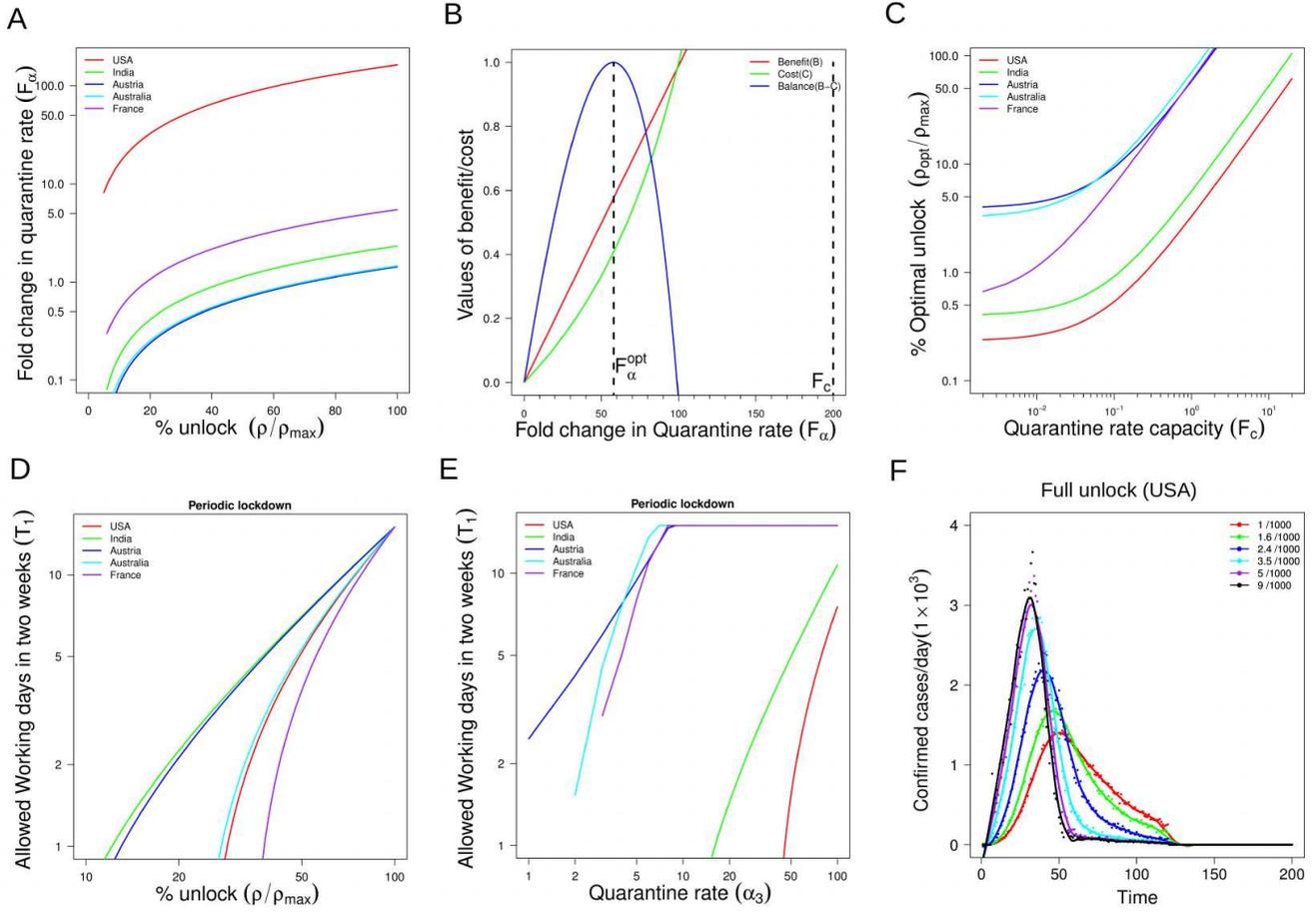
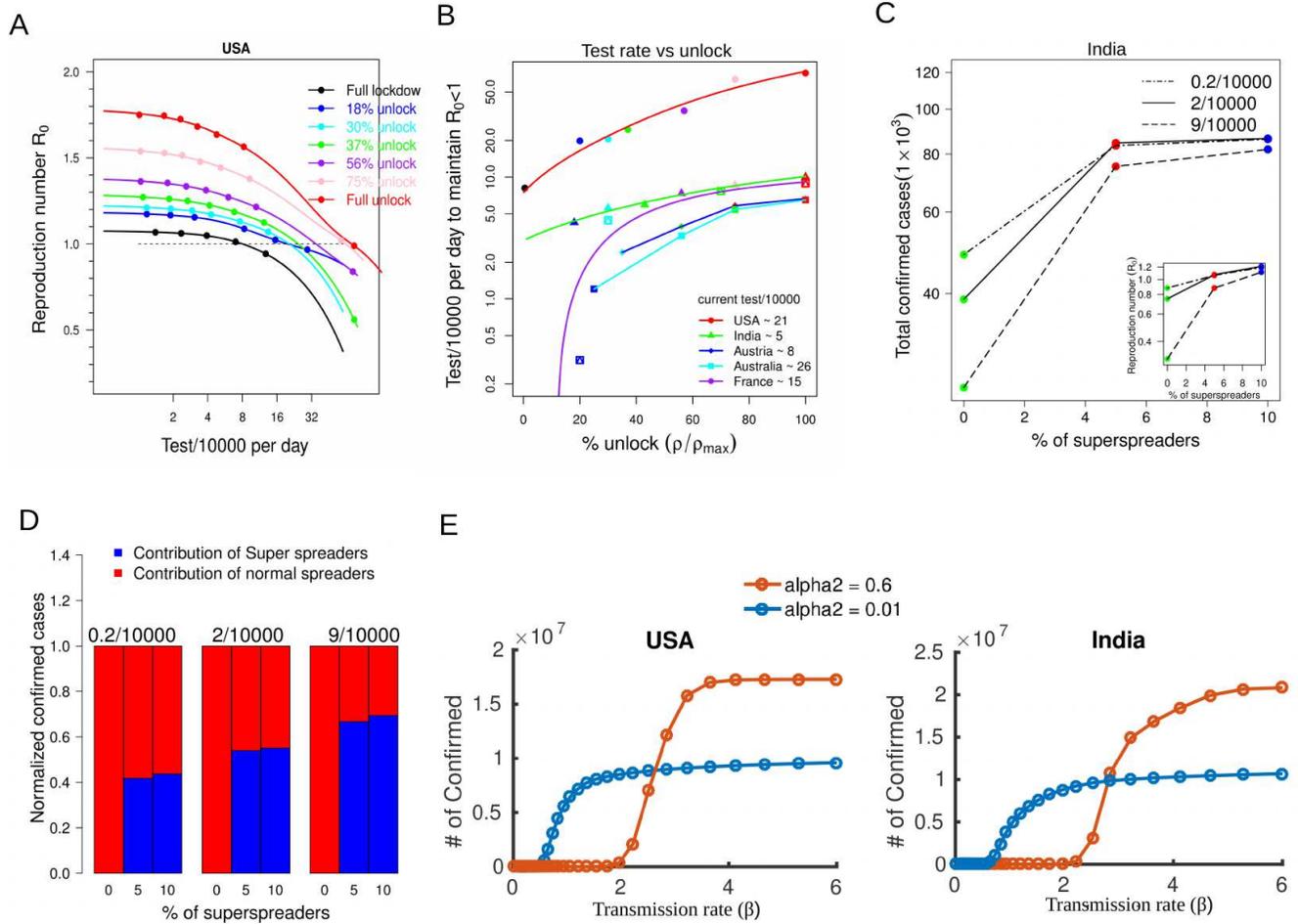


Figure 5



Figures

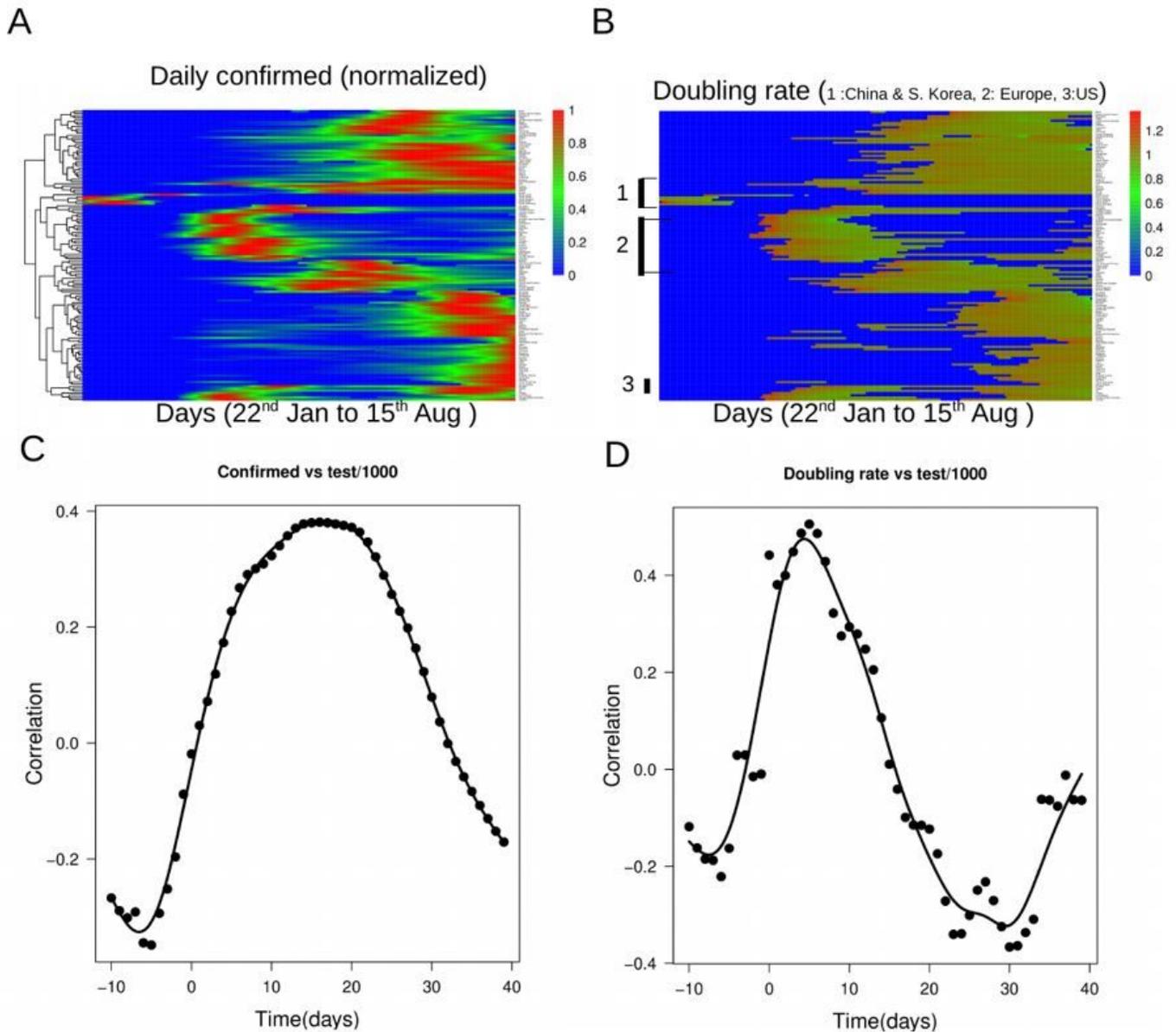


Figure 1

The Covid-19 outbreak in different countries and its relation with the testing rate. (A) The heat map displays the clustered dynamics for around 100 different countries for the daily cases normalized to the maximum for each country (Days correspond to 22nd January, 2020 and to 15th August, 2020). The dendrogram is based on hierarchical clustering of the time traces (B) The heat map represents the doubling rate as a function of time clustered according to the dynamics of the time traces as in (A). The results are shown in exponential scale. (C) The Pearson correlation coefficient of test rate with the daily confirmed cases at different time points along the dynamics are shown. The time traces are aligned w.r.t the maximum doubling rates. 0 on the x-axis corresponds to the maximum doubling rate for all the

countries. (D) The Pearson correlation coefficient of test rate with the doubling rate at different time points along the dynamics aligned as described in (C).

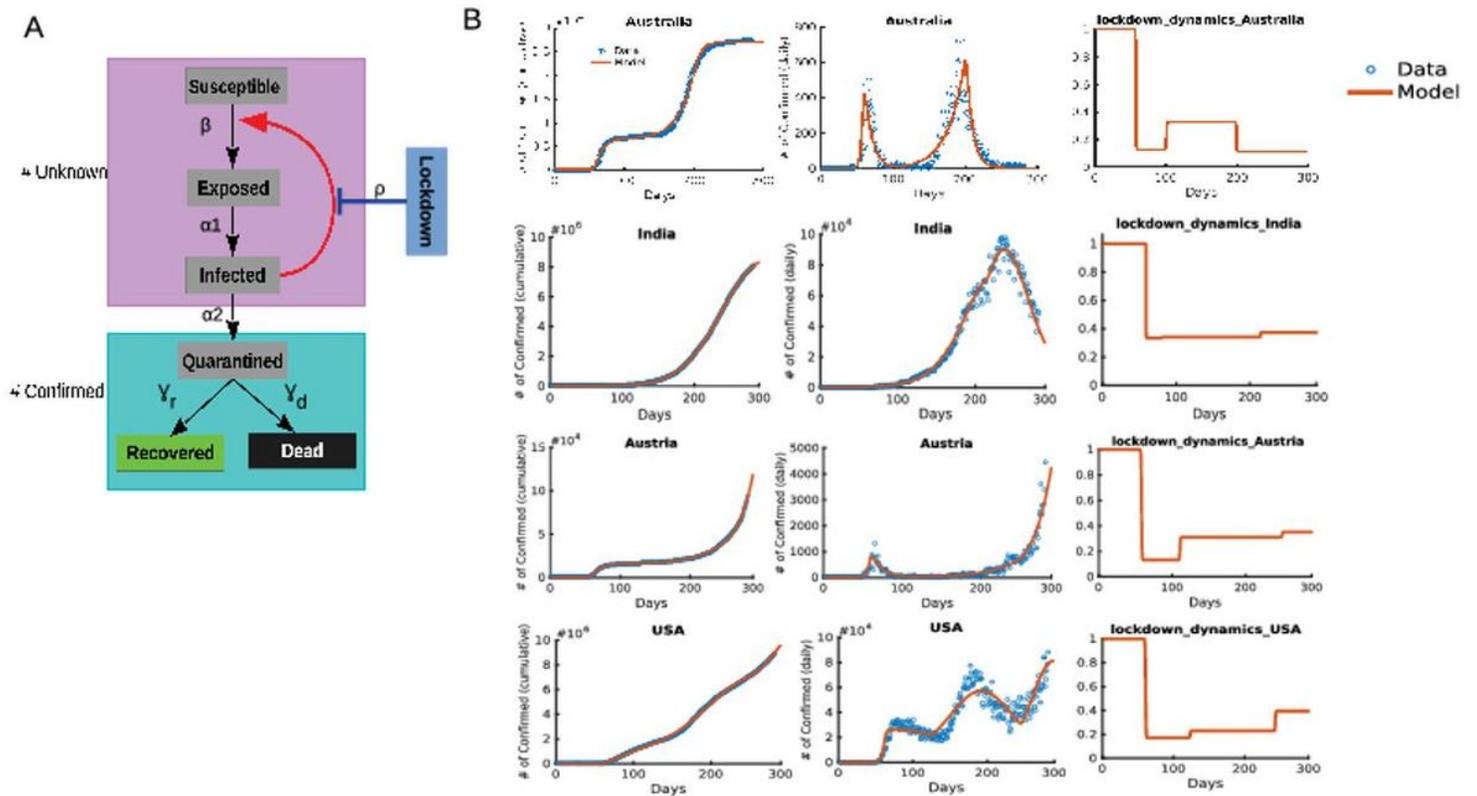


Figure 2

The description of the SEIQR model and calibration for different countries. (A) Schematic representation of the SEIQR model which contains susceptible, exposed, infected, quarantined, recovered and dead compartments. The lockdown is implemented through a sigmoid function as shown in the methods section. The quarantined, recovered and dead cases together comprise the confirmed cases. The arrow from infected to susceptible represents the positive feedback that fuels the infection spread in the population which is negatively regulated by the lockdown, as indicated (B) The SEIQR models fit the data for 4 representative countries as indicated for the cumulative confirmed and daily confirmed cases. The number of days in X axis corresponds to the time course data available in JHU CSSE [41] where 0 corresponds to 22nd January, 2020 and the end time point corresponds to 30th October, 2020. The lockdown function shows the extent of lockdown with time where 0 = full lockdown and 1 = no lockdown.

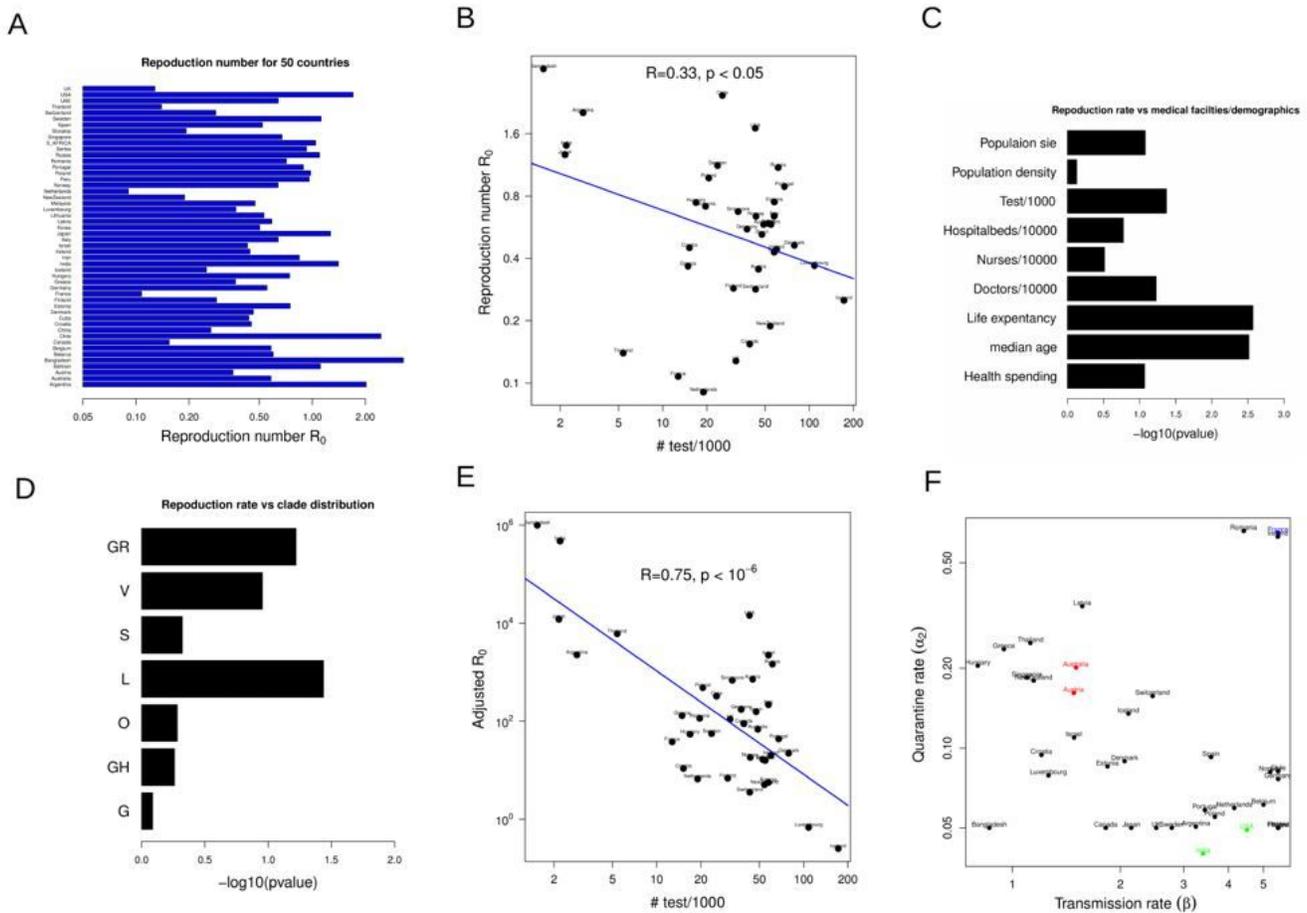


Figure 3

The reproduction number (R_0) displays correlation with the testing rate across countries. (A) The barplot shows R_0 for the fitted 50 countries after lockdown. (B) The scatter plot describes correlation between the test rate and R_0 after lockdown for the 39 countries with at least 10 complete genome sequences as of end of June described in the main text. (C) The statistical significances of the dependence of R_0 on different demographic and medical facilities factors are shown based on linear regression based on the same 39 countries (D) The same statistical significance of dependence of R_0 on the frequencies of the clades (G,GR,GH,L,O,S,V) are shown. (E) The scatter plot indicates the correlation between the adjusted R_0 values and the test rates. (F) The values of the fitted transmission rates and quarantine rates for all the 50 countries are plotted and the representative countries for consideration of the agent based model simulation are shown in colors red, green and blue.

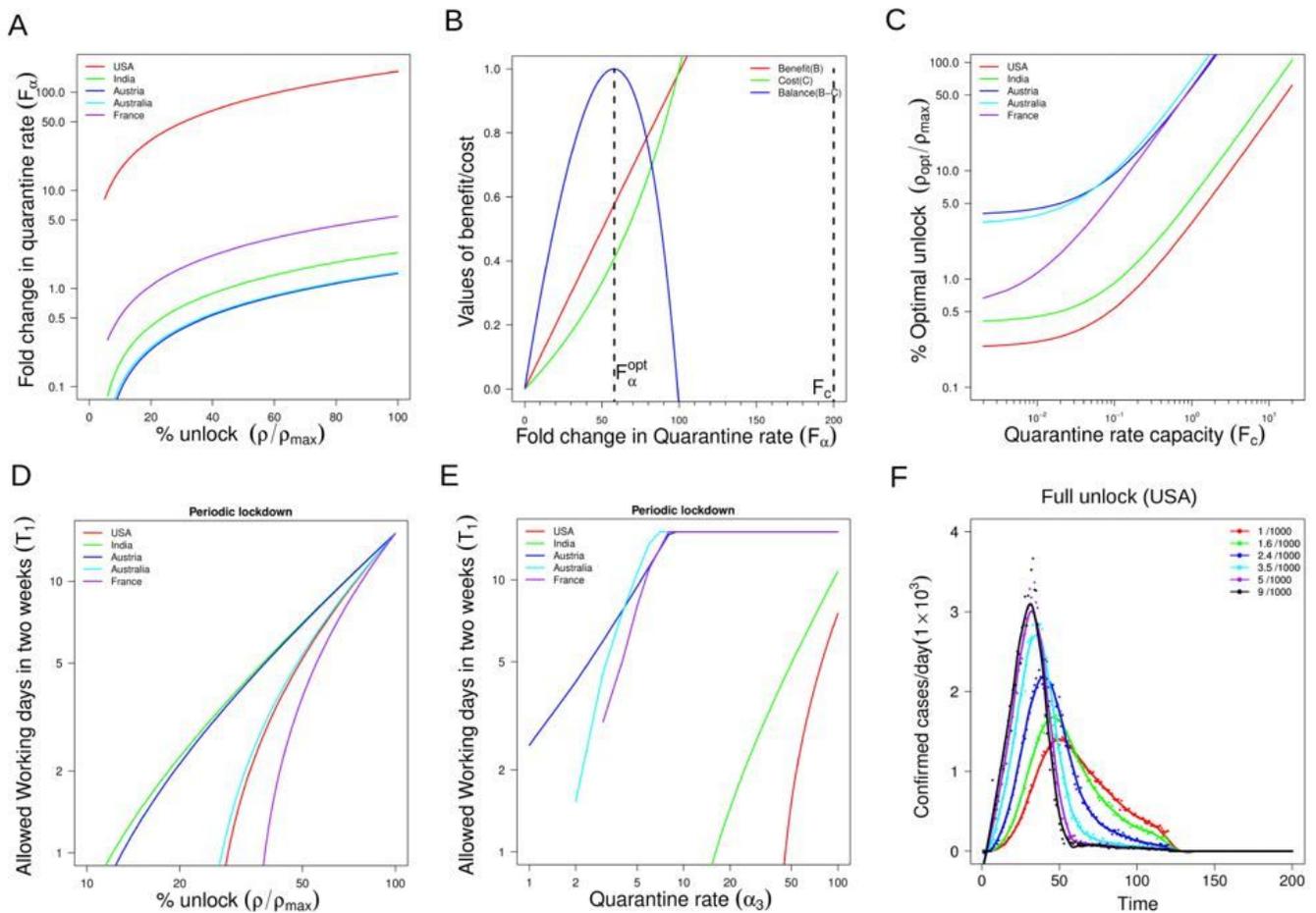


Figure 4

The optimization of trade-off between testing rate and extent of unlock. (A) The relationship between the extent of unlock and the corresponding quarantine rate required to keep R_0 values equal to one for five countries as indicated. The y-axis is plotted in log-scale. (B) The cost and benefit curves as a function of quarantine fold change for the countries. The cost benefit is optimized at a value F_{α}^{opt} for a maximum capacity F_c of the fold change in quarantine rate. (C) The optimum extent of unlock calculated from F_{α}^{opt} as a function of the maximum quarantine capacity F_c . Both the axes are plotted in log-scale. (D) The implementation of the extent of unlock as a function of the periodic unlock for the different countries are indicated. Number of working days represents the working day cycle in 15 days time. (E) The quarantine required to be increased by testing to keep the R_0 value below one and the corresponding the working days allowed. (F) The time traces of the confirmed daily cases of agent based simulation based on parameter values with full unlock for the USA at different daily tests/10000 as indicated.

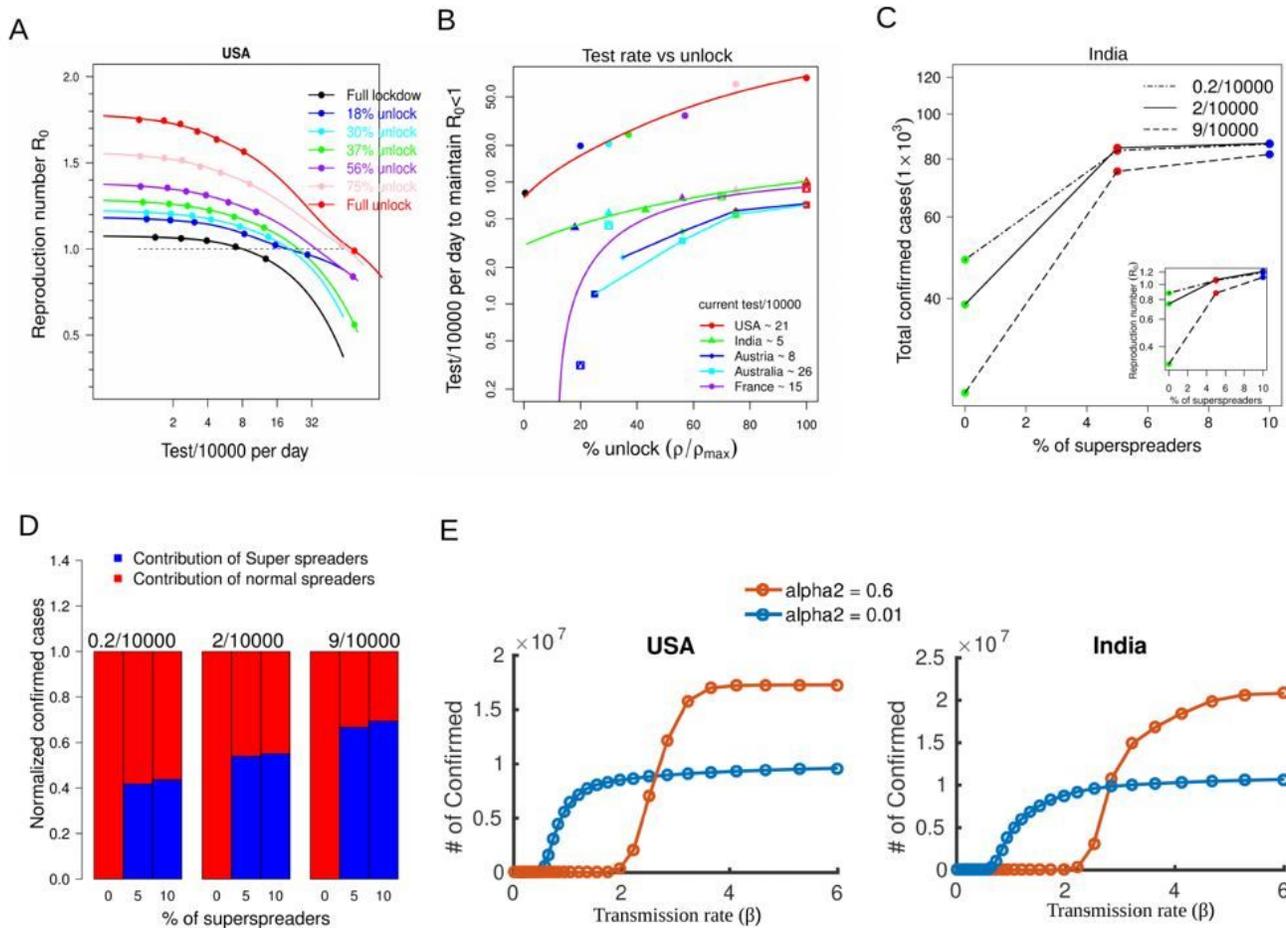


Figure 5

The testing rate and fraction of superspreaders in the population influence the infection rate simulated by the agent based model (A) The R_0 values based on the agent based model with parameter values of the USA as a function of test/10000 for different extents of unlock as indicated . The other four countries are shown in supplementary figure S4. The dashed horizontal line exhibits the $R_0=1$ line, intersection of the line with the curves quantifies the test/10000 required to maintain $R_0=1$. (B) The allowed extent of unlock and the corresponding test/10000 required to keep the R_0 value less than one calculated from the agent based simulation for the different counties with indicated current test rate. (C) The total confirmed cases as a function of percentage of superspreaders simulated for parameter values of India at different testing rates as indicated. (The inset shows the corresponding R_0 values at different percentages of superspreaders). (D) The barplot indicates the contribution of superspreaders and normal spreaders over the total number of confirmed cases as a function of the percentage of superspreaders at 3 different testing rates 0.2/10000, 2/10000 and 9/10000, as indicated. (E) The relationship between the transmission rate and the total confirmed cases at two different values of quarantine rate as indicated for two representative countries, USA and India.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SI.pdf](#)