CALCULATING THE EFFECTIVE REPRODUCTION NUMBER FOR COVID-19 USING A NEW PROCESS FOR VARIOUS COUNTRIES

A TECHNICAL REPORT

Patrick Grice
CloseAssociate (plc)
Wellington, New Zealand
Patrick.Grice@closeassociate.com

Stephen Grice, Ph.D.
CloseAssociate (plc)
Wellington, New Zealand
Stephen.Grice@closeassociate.com

Prof. Richard S. Laugesen
Department of Mathematics
University of Illinois at Urbana-Champaign, USA
Laugesen@illinois.edu

May 26, 2020

ABSTRACT

A new process for estimating R_{eff} for COVID-19 is developed, combining a deterministic SIR formula for calculating R_{eff} from positive test data with a statistical bootstrapping method for generating confidence intervals. This SIR+B process is used to estimate R_{eff} for a selection of countries, and it produces results that compare favourably with other estimation methods. The paper suggests the SIR+B approach offers significant potential to inform public health policy.

1 Introduction

SIR modelling is effective at predicting disease spreading through populations [1]. The dominant parameter in all SIR models is the effective reproduction number R_{eff}. This effective reproduction number is a function of time t, since it changes as the number of Susceptible individuals reduces during an epidemic, and as communities implement social distancing, quarantining and contact tracing measures. Estimating R_{eff} is critical to understanding disease progression in a population, and hence is of intense interest to public health experts and policy officials.

Disease control requires keeping the effective reproduction number $R_{\text{eff}}(t)$ below the critical value of 1.

This technical report develops a new "SIR plus bootstrapping" (SIR+B) process for estimating $R_{\text{eff}}(t)$ for a range of countries. It provides public officials with a rapid and effective tool for knowing whether the reproduction number is increasing or decreasing over time, and when it is less than or greater than 1.

The SIR+B process is simpler than statistical approaches to estimating R_{eff} such as the "EpiEstim" optimization approach of Cori et al. [2] and Thompson et al. [3] and references therein. The EpiForecasts website [4] generates statistical estimates of R_{eff} for many countries and regions around the world. It assumes the reproduction number has mean 2.6 and standard deviation 2.0, following a gamma distribution [6]. This hard-coded mean of 2.6 presumably explains why EpiForecasts' estimates of the COVID-19 reproduction number consistently start out around 2.0–2.5, regardless of the country or region being modeled.

The deterministic part of the process in this paper computes R_{eff} explicitly from the SIR model by formula (9), rather than arriving at the reproduction number statistically, and then we apply bootstrapping to generate confidence intervals. We will graphically compare our results for $R_{\text{eff}}(t)$ with EpiForecasts, in figures in the next section.

The SIR+B process produces outputs that are directly relevant to public policy. Results are displayed graphically for several countries and regions in Section 2. This is followed by a conclusion in Section 3. The technical foundations of the new process are summarized in the Appendices, with full details to come soon [5].

2 Results on effective reproduction number

The estimates of R_{eff} from our SIR+B process are presented below for a selection of countries. Results are correlated with the implementation of social distancing, testing and tracing policies in many countries. EpiEstim's graphs of the reproduction number are plotted also, for purposes of comparison.

2.1 Six countries with large numbers of cases — United States, United Kingdom, Russia, Peru, India, Brazil

Country-specific descriptions are given in the captions of each figure. In each country, the disease spread widely ($R_{\text{eff}} \gg 1$) before effective mitigations were in place, and the mitigations implemented have not been strong enough to prevent substantial numbers of new infections.

The first country in this series is the **United States**, see Figure 1. There the federal government plays a secondary policy role in response to COVID-19, with the governor in each State leading mitigation policy and responses. Each State has put in place a different set of restrictions, lockdowns and social distancing measures. Throughout March and early April, several States, cities, and county governments imposed "stay at home" orders on their populations to stem the spread of the virus. New York was hardest hit, with hospitalisation peaking in the first week of April. Other States followed. The relatively delayed onset of mitigation efforts has resulted in just a slow decrease in the effective reproduction number, with $R_{\text{eff}} \geq 5.0$ on or about 13 March, decreasing to $R_{\text{eff}} \approx 1.0$ on or about 23 April. Overall, the United States has had $R_{\text{eff}} \approx 1.0$ from that time, although this average is known to mask significant regional differences.

The second country is the **United Kingdom**, see Figure 2. On 22 January the United Kingdom government announced a policy to contain, delay, research and mitigate the disease, with an apparent strategy to obtain "herd immunity" over time while protecting the National Health Service. Britain remained open, resisting the kind of lockdowns imposed elsewhere in Europe [13]. On the 23 March, prime minister Johnson changed course and announced tightening new mitigations. These delayed mitigation efforts in the U.K. are reflected in a slow decrease in the effective reproduction number, from $R_{\text{eff}} \geq 4.0$ on 13 March to $R_{\text{eff}} \approx 1$ by 15 April. It has continued to decrease, and we estimate $R_{\text{eff}} \approx 0.7$ by 16 May.

The third country is **Russia**, see Figure 3. Russia and its districts and surrounding republics announced a series of lockdowns and social distancing measures beginning in the period 28 March to 1 April. As a result, the effective reproduction number has been decreasing from $R_{\text{eff}} \geq 3.0$ on 30 March to $R_{\text{eff}} \approx 1.1$ on 16 May. The EpiEstim method consistently gives lower estimates of R_{eff} than our process does.

The fourth country in this series is **Peru**, see Figure 4. The government of Peru has issued a series of escalating quarantining and social measures starting 15 March. Compliance with the quarantining and social distancing measures has been uneven. R_{eff} has oscillated and decreased over this period of time. Recently the effective reproduction number has stabilised at a lower value $R_{\text{eff}} \approx 1.5$. (Smaller oscillations are also visible but are underestimated in the EpiForecast results.) The epidemic will continue growing exponentially in Peru unless additional measures are taken.

The fifth country is **India**, see Figure 5. On 24 March the Indian prime minister Narendra Modi announced a complete nationwide lockdown for 21 Days [23] until 14 April. On 14 April, the prime minister extended the nationwide lockdown to 3 May, with a conditional relaxation from 20 April for the areas that had been able to contain the spread. R_{eff} decreased from ≈ 3.2 to 1.8 from the beginning to the end of the lockdown period. A subsequent increase in R_{eff} with a maximum around 1 May may relate to the conditional relaxation. These effects are not displayed in the EpiForecasts results, which we believe demonstrates the higher resolution of the new process.

The last country in this series is **Brazil**, see Figure 6. The new process in this paper estimates R_{eff} to be significantly higher than the EpiForecast estimates from 4 March to 16 May. The new process estimates $R_{\text{eff}} \approx 1.8$ from about 7 April to the present. We believe the EpiForecast method significantly underestimates R_{eff}. The government of Brazil has been in conflict with state governors in implementing COVID-19 social distancing policies [8]. On 6 May Brazil reported about 7,000 positive tests, and on 15 May the number of positive tests doubled to about 14,000. This rapid rise in cases suggests that $R_{\text{eff}} \geq 2.0$, giving confidence that the new process is more informative than EpiForecasts, for Brazil. The rapid rise in new cases suggests that R_{eff} is substantially larger than 1, and that the new process is providing more accurate estimates for Brazil than the EpiEstim method.

2.2 Four countries with small numbers of cases — New Zealand, South Korea, Australia and Iceland

In the second series of figures we examine four countries that have quickly and successfully reduced R_{eff}. These countries were quick to respond in the beginning stages of the epidemic, and rapidly implemented mitigations strong enough to prevent exponential growth.

The first country in this series is **New Zealand**, see Figure 7. Of all the countries in this series, New Zealand has employed the most strict social distancing policy – a full society lockdown implemented from 26 March. This had the effect of reducing R_{eff} more rapidly than any other country examined. Within 15 days of the lockdown, community transmission of the disease had stopped, and $R_{\text{eff}} \approx 0.0$ from 10 April, meeting the World Health Organisation definition of "elimination" of the disease. After that date, R_{eff} is no longer knowable, because there are so few new cases that the SIR model on which we base the initial deterministic computation of R_{eff} is no longer applicable.

New Zealand mitigated the epidemic largely through strict social distancing policies.

The second country is **South Korea**, see Figure 8. South Korea has implemented a series of strict policies based on aggressive testing, tracing and isolation of exposed and infected individuals. These policies have resulted in a rapid decrease in R_{eff} dating from 26 February, and have continued to ensure that $R_{\text{eff}} \leq 1.0$ until 2 May. After a sustained period with daily reported cases in the country below 20, a new cluster emerged in central Seoul. A 29-year-old infected person from Yongin was found to have visited at least five night clubs in Itaewon during the late night hours of May 1 and the subsequent early morning hours of May 2. This new cluster resulted in 79 new daily cases being reported by 11 May. This cluster can be seen in the peak in R_{eff} of 2.0 occurring on or about 10 May. The new cluster that occurred on 2 May has been successfully controlled with $R_{\text{eff}} < 1.0$ by about 13 May, by using contact tracing and isolation policies.

The third country in the series is **Australia**, see Figure 9. Australia has used the federal system to implement a combination of aggressive testing, quarantining and social distancing. These policies resulted in a fast decrease in R_{eff}. The number of new cases initially grew sharply, then levelled out at about 350 per day around 22 March, and started falling at the beginning of April to under 20 cases per day by the end of the month [15]. As of 18 May 2020, 7,060 cases and 99 deaths had been reported in Australia, with the highest number of cases being in New South Wales, with 3,076. The effective reproduction number decreased rapidly from $R_{\text{eff}} \geq 5.0$ on or about 13 March to $R_{\text{eff}} \approx 0.3$ on or about 1 April, and $R_{\text{eff}} < 1$ since that time. A recent increase in R_{eff} might signal the appearance of a new cluster at the time of writing. We believe the EpiEstim method has overestimated R_{eff} in Australia. Also, note it does not reach close enough to the present day to signal the possible new cluster.

Australia mitigated the epidemic with a combination of aggressive testing, quarantining and social distance controls that were less restrictive than in New Zealand.

The last country considered is **Iceland**, see Figure 10. Iceland has a small population but has acted early with aggressive and comprehensive testing, isolation and quarantining, and mild social distance policies. On 13 March, it was announced that public gatherings of more than 100 would be banned and universities and secondaries schools closed for four weeks. By 24 March, a nationwide ban on public assemblies over 20 took effect. All swimming pools, museums, libraries and bars closed, as did any businesses requiring a proximity of less than 2 metres. The impact on aggressive testing, quarantining and mild social distancing can been seen in the decrease in the effective reproduction number with $R_{\text{eff}} \geq 3.0$ on or about 13 March and decreasing to $R_{\text{eff}} \approx 0.2$, before being eliminated by 15 April. After this time there is no observable community transmission and R_{eff} cannot be measured. Given the small population of Iceland, the EpiEstim method may not be reflective of the spread of the epidemic there.

2.3 Three similar countries with dissimilar outcomes — Italy, Germany, France

The last series of countries we examine are European nations that are geographically close and with large populations, but with very different outcomes.

The first country in this set is **Italy**, see Figure 11. Italy was one of the first European states to experience COVID-19 infections, with its first confirmed cases being a Chinese couple, originally from Wuhan, who arrived on 23 January via Milan Malpensa Airport [16]. The Italian government response was uneven, and the epidemic grew to the point of overwhelming the health system in the north. Lombardy and other regions were in crisis by early March. On 10 March, prime minister Conte increased the quarantine lockdown to cover all of Italy, including travel restrictions and a ban on public gatherings. On 21 March, Conte announced further restrictions with a nationwide lockdown and halting of all non-essential production, industries and businesses. The slow and varied response of the government is reflected in the slow decrease in the effective production number, with $R_{\text{eff}} \geq 4.0$ on 28 February, decreasing to $R_{\text{eff}} \leq 1.0$ on or about 30 March. Since that time $R_{\text{eff}} \leq 1.0$. The EpiForecast estimates understate the reduction in R_{eff}, in our opinion.

The second country is **Germany**, see Figure 12. The first confirmed case of COVID-19 in Germany occurred on 27 January [17]. The government response was enacted at the state level, by order of federal health minister Jens Spahn on 26 February [18]. Restrictions were limited to quarantining infected cases, limited travel restrictions (notably to areas with high numbers of cases) and ad hoc cancellations of large events; however, there were no national rules. On 14 March, a new Cabinet Committee was set up and over the next week country-wide restrictions were put in place,

including closing schools and religious services, shops, and businesses. Although the overall trend of the effective reproduction number R_{eff} for Germany looks similar to that of Italy, Germany achieved $R_{\text{eff}} \leq 1.0$ much sooner. The value $R_{\text{eff}} \geq 4.0$ on 13 March decreased to $R_{\text{eff}} \leq 1.0$ on or around 1 April, achieving this reduction in about half the time of Italy, demonstrating that coordinated regional responses can be more effective than slow national approaches. Since that time, the value has stabilized with $R_{\text{eff}} < 1.0$.

The third country in this set is **France**, see Figure 13. The first confirmed case of COVID-19 in Europe was on 24 January in France [19]. (Earlier cases from December have since been confirmed as COVID-19). After a number of clusters emerged across the country, the French government enacted restrictions on the size of gatherings, reducing the maximum allowed from 5,000 on 5 March to 100 on 14 March. Municipal elections occurred on 15 March and the official lockdown period began on 16 March with wider restrictions on public services, institutions, and travelling. Since then, incremental increases and extensions to lockdown restrictions were implemented. From 11 May, France began to ease lockdown restrictions in incremental steps. The incremental response of the government is reflected in the slow but steady decrease in the effective reproduction number, with $R_{\text{eff}} \geq 4.0$ on 26 February, decreasing to $R_{\text{eff}} \leq 1.0$ for the first time on or about 31 March.

For each of these large countries, Italy, Germany and France, the reduction in R_{eff} is more pronounced using our process than with the EpiForecasts method.

2.4 Sensitivity analysis

The SIR+B process depends only on the daily case data, the recovery parameter γ, and the testing parameter c (see the Appendices). Issues relating to the accuracy of the daily case data are described in a separate technical report [5]. All that remains is to understand the sensitivity of the method to different choices in γ and c.

We have chosen $\gamma = 1/15$ for our analyses, which seems a sensible value on the basis that individuals take an average of about 15 days from initial infection to recovery.

We are not aware of any firm evidence for the value of the testing rate c. This value will likely differ from country to country, as governments respond to the epidemic with varying degrees of urgency.

To address sensitivity of the results to the parameter values, we examine Germany and the U.S. in Figures 14 and 15, and find that changes in the testing rate are not influential on the results for R_{eff}. Allowing γ to range from $1/10$ to $1/20$ generally preserves the nature of the results too.

3 Conclusion

A SIR+B process for estimating the effective reproduction number from daily case data has been introduced and applied to a variety of countries. The process compares favourably with statistical methods in use, because:

- it is faster and is numerically stable,
- it can provide early warning before the health system is overwhelmed,
- it provides additional information on new clusters and their suppression, as demonstrated recently in South Korea,
- it gives a more transparent estimation of the errors,
- the new SIR+B process requires only two input parameters, which are the testing rate and the recovery rate (see the appendices).

Thus the paper demonstrates the potential of the SIR+B process to inform public health policy in the immediate future.

[Figures begin on the next page.]

Data source and interpretation of figures. All graphs below use the daily case data from the European Centre for Disease Prevention and Control [7]. This is also true for the EpiForecast plots. The deterministic R_{eff} calculated from the raw data through an SIR model are shown as black squares. The deterministic method combined with the bootstrapping process in this paper generates the 50% confidence intervals shown in dark grey, and 95% confidence intervals in light grey. The width of the confidence intervals is proportional to the uncertainty in the number of daily new cases and to the number of daily new cases. The critical value $R_{\text{eff}} = 1$ is shown as a dashed red line. The EpiForecast results are shown in blue (50% credible limit) and light blue (90% credible limit), for purposes of comparison with the new process.

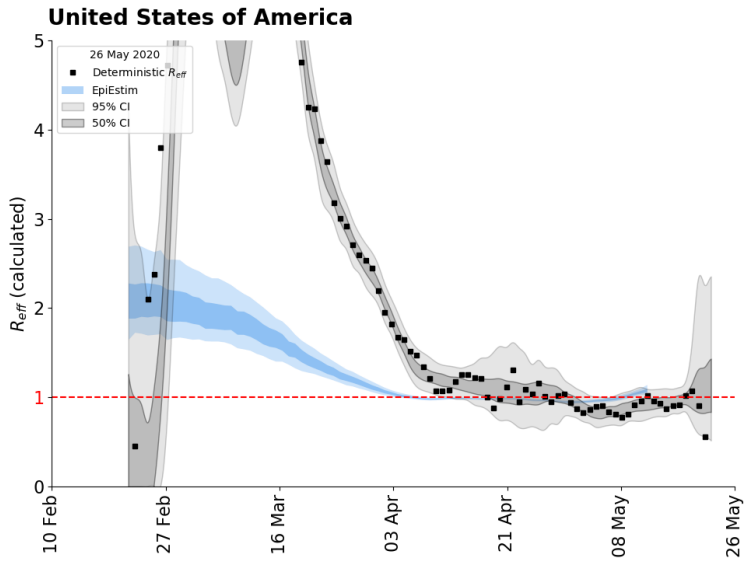

Figure 1: **United States R_{eff} results compared with EpiForecasts.**
The federal government of the United States plays a secondary policy role in response to COVID-19, with the governors of each State leading mitigation policy and response for their State. Each State has had a different set of restrictions, lockdowns and social distancing measures. Throughout March and early April, several States, cities, and county governments imposed "stay at home" orders to stem the spread of the virus. New York was hardest hit, with hospitalisations peaking in the first week of April. Other States followed. The slow national response resulted in only a slow decrease in the effective reproduction number, with $R_{\text{eff}} \geq 5.0$ on or about 13 March and decreasing to $R_{\text{eff}} \approx 1.0$ on or about 23 April. Overall the United States has stayed $R_{\text{eff}} \approx 1.0$ from that time. The R_{eff} values estimated by the new process in this paper are consistently higher than estimates by EpiForecasts

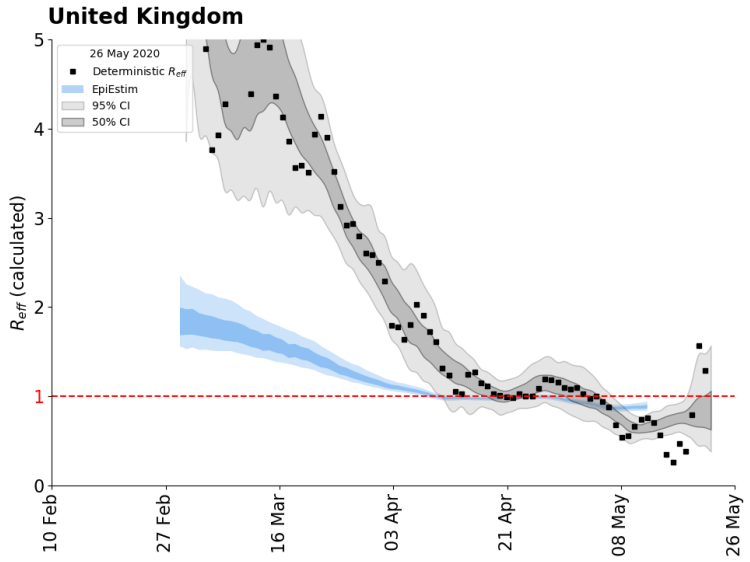

Figure 2: **United Kingdom R_{eff} results compared with EpiForecasts.**
The R_{eff} values estimated by the new process in this paper are higher than estimates by EpiForecasts.
On the 22 January the United Kingdom government announced a policy to contain, delay, research and mitigate the disease to obtain "herd immunity" while protecting the National Health Service. Britain remained open, resisting lockdowns seen elsewhere in Europe [13]. Prime minister Johnson changed strategy and on 23 March announced tightening new mitigations. The slow response to the epidemic is reflected in the slow decrease in the effective reproduction number, going from $R_{\text{eff}} \geq 4.0$ on 13 March to $R_{\text{eff}} \approx 1.0$ by 15 April. It has continued to decrease, and $R_{\text{eff}} \approx 0.7$ by 16 May.

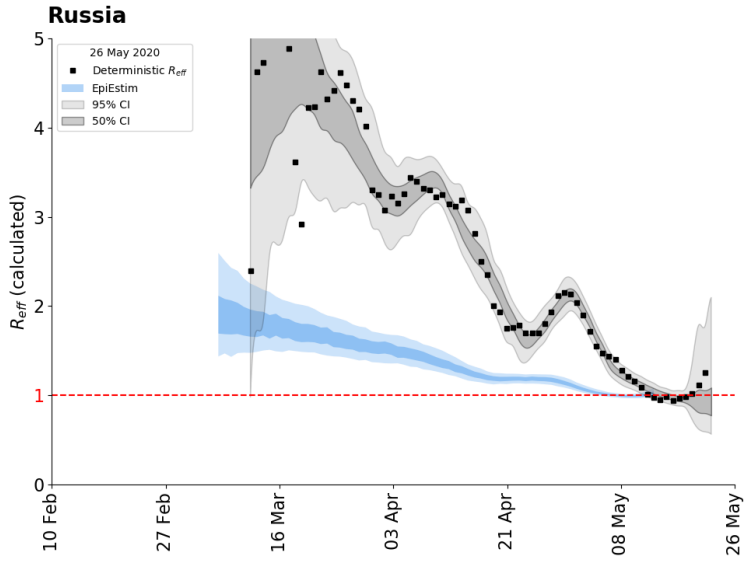

Figure 3: **Russia R_{eff} results compared with EpiForecasts.**
The R_{eff} values estimated by the new process in this paper are higher than estimates by EpiForecasts.
Russia, its districts and surrounding republics announced a series of lockdowns and social distancing measures beginning around 28 March to 1 April, reducing the effective reproduction number.

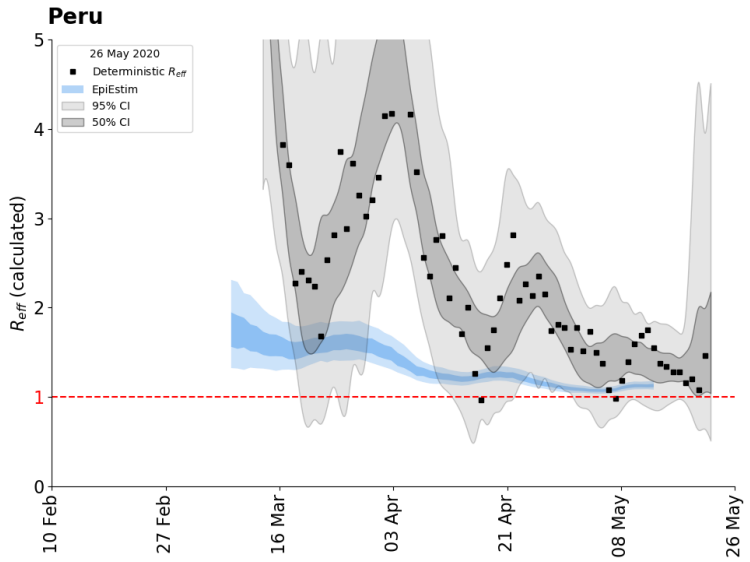

Figure 4: **Peru R_{eff} results compared with EpiForecasts.**
The R_{eff} values estimated by the new process in this paper are higher than estimates by EpiForecasts.
The government of Peru has issued a series of escalating quarantining and social measures starting 15 March. Compliance with the quarantining and social distancing measures has been uneven. R_{eff} has oscillated and decreased over this period of time. Recently the effective reproduction number has stabilised at a lower value $R_{\text{eff}} \approx 1.5$.

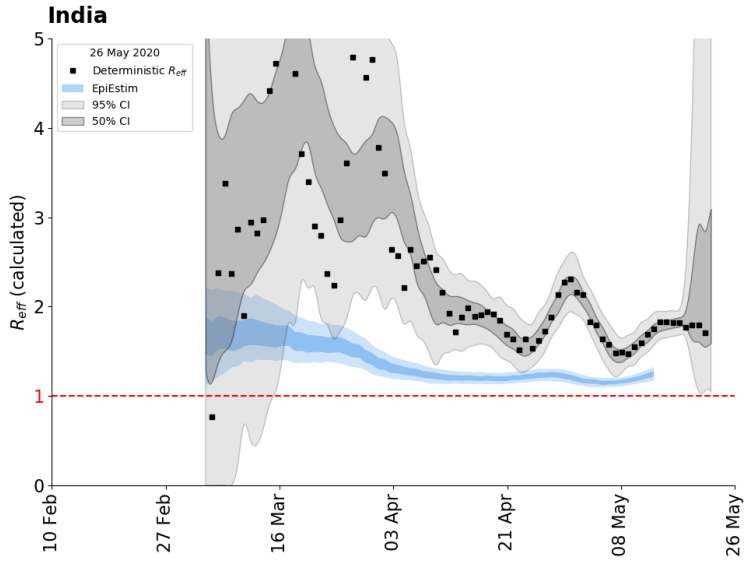

Figure 5: **India R_{eff} results compared with EpiForecasts.**
On 24 March the Indian prime minister Narendra Modi announced [9] a complete nationwide lockdown for twenty one days until 14 April. On 14 April, he extended the nationwide lockdown till 3 May, with a conditional relaxation from 20 April for the areas that had been able to contain the spread. From the beginning to the end of the lockdown period, R_{eff} decreased from ≈ 3.2 to 1.8. There is an increase in R_{eff} with a maximum at or about 1 May which may relate to the conditional relaxation. These effects are not displayed in the EpiForecasts results, which also are substantially lower than our estimates of R_{eff} for India.

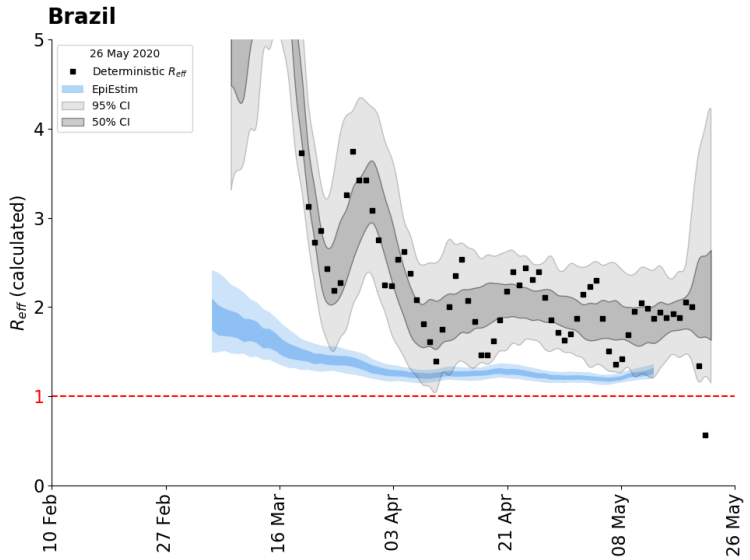

Figure 6: **Brazil R_{eff} results compared with EpiForecasts.**
The R_{eff} values estimated by the new process in this paper are higher than the estimates by EpiForecasts. The most recent estimate is $R_{\text{eff}} \approx 1.8$, and is constant from about 7 April to present. The government of Brazil has been in conflict with state governors in implementing COVID-19 social distancing policies [8]. On 6 May Brazil reported about 7,000 positive tests, and on 15 May the number of positive tests doubled to about 14,000. This rapid rise in cases suggests that $R_{\text{eff}} \geq 2.0$, giving confidence that the new process is more informative than EpiForecasts, for Brazil.

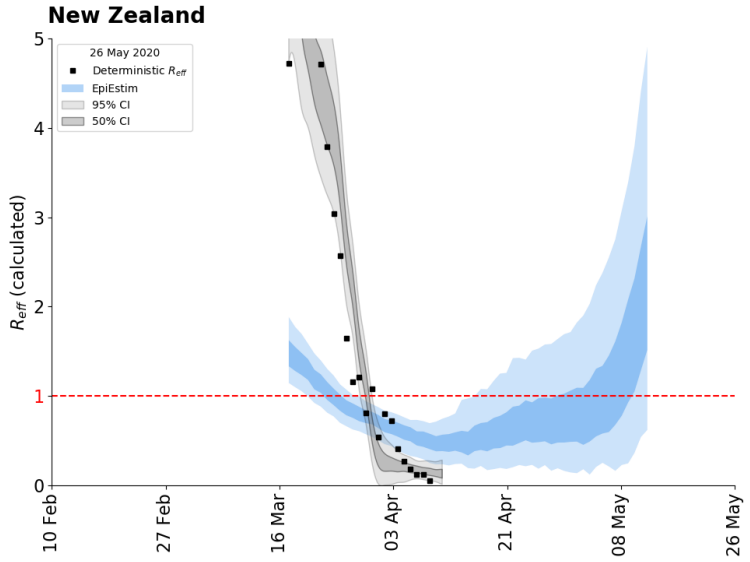

Figure 7: **New Zealand R_{eff} results compared with EpiForecasts.**
On 21 March, New Zealand prime minister Jacinda Ardern introduced a country-wide alert level system to deal with the coronavirus outbreak [10]. The alert level was set to 2 immediately. On 23 March, Ardern raised the alert level to 3 and announced the closure of all schools. All sports matches and events as well as non-essential services were required to close in 48 hours. Essential services such as supermarkets, petrol stations, and health services remained open. The social distancing policies had an immediate effect on R_{eff}. Leading into the lockdown $R_{\text{eff}} \geq 4.0$, although this estimate is possibly artificially high due to New Zealanders returning home with the disease. By the second week of April community transmission had stopped, and $R_{\text{eff}} \approx 0$ on or about 10 April. After this date R_{eff} is undefined as the signal of the epidemic is lost. Among all countries in the world, the New Zealand data shows the most rapid decline in R_{eff} and New Zealand is one of only a few countries to achieve the World Health Organisation definition of elimination of COVID-19 [11, 12].

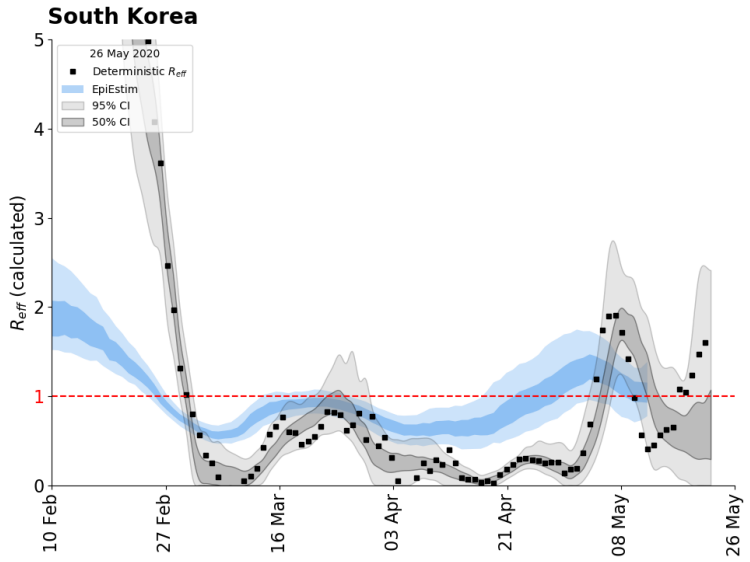

Figure 8: **South Korea R_{eff} results compared with EpiForecasts.**
South Korea introduced what was considered one of the largest and best-organised epidemic control programs in the world. Different measures have been taken to screen the mass population for the virus, and isolate any infected people as well as trace and quarantine contacts of the infected, but did not implement a "stay at home" order [14]. The rapid and extensive tests taken by South Korea have been judged successful in limiting the spread of the outbreak, without using "stay at home" orders. The rapid and aggressive testing policy lead to a rapid decrease in the effective reproduction number, with $R_{\text{eff}} \geq 5.0$ on or about 26 February and decreasing to $R_{\text{eff}} \approx 0.0$ by 8 March. $R_{\text{eff}} \leq 1.0$ from this date. After a sustained period of reported daily cases in the country being below 20, a new cluster emerged in central Seoul. A patient from Yongin was found to have visited at least five nightclubs in Itaewon during the late night hours of 1 May and early morning hours of 2 May. This new cluster resulted in 79 new daily cases being reported by 11 May. This cluster can be seen in the peak in R_{eff} of 2.0 occurring on or about 10 May. The general features of our R_{eff} plot are visible also in the EpiForecast graph, but are understated and not as detailed there. For example the new process shows that the cluster that occurred on 2 May was controlled with $R_{\text{eff}} < 1.0$ by about 13 May. This specific feature is not evident in the EpiForecast nowcast data.

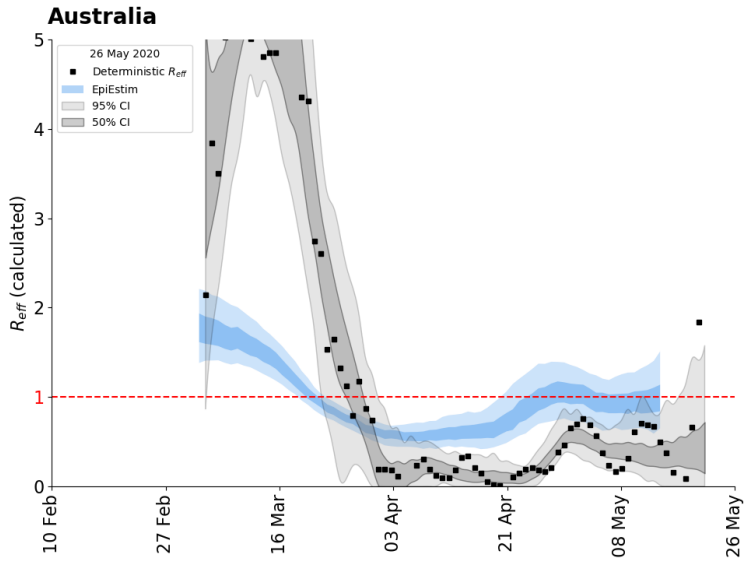

Figure 9: **Australia R_{eff} results compared with EpiForecasts.**
On 20 March, Australia closed its borders to all non-residents and non-Australians. The Australian Government imposed the ban in coordination with New Zealand. This has been accompanied in Australian States with milder social distancing rules than New Zealand has implemented. The number of new cases in Australia initially grew sharply, then levelled out at about 350 per day around 22 March, and started falling at the beginning of April to under 20 cases per day by the end of the month [15]. As of 18 May 2020, 7,060 cases and 99 deaths had been reported in Australia, with the highest number of cases being in New South Wales, with 3,076. The effective reproduction number decreased rapidly from $R_{\text{eff}} \geq 5.0$ on or about 13 March to $R_{\text{eff}} \approx 0.3$ on or about 1 April. $R_{\text{eff}} < 1$ since that time. A recent increase in R_{eff} could signal the appearance of a new cluster at the time of writing.

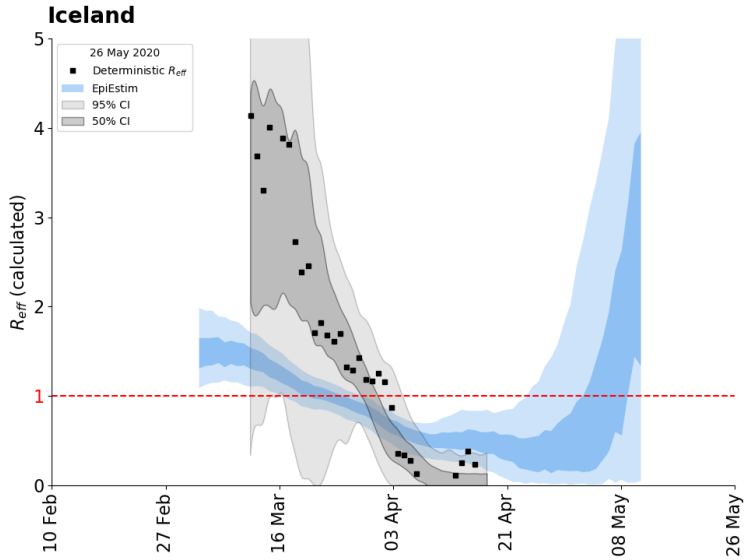

Figure 10: **Iceland R_{eff} results compared with EpiForecasts.**
The response to the pandemic by Icelandic health authorities has focused on early detection and contact tracing, and social distancing measures such as a ban on assemblies of more than 20 persons. On 13 March, it was announced that public gatherings of more than 100 would be banned and universities and secondary schools closed for four weeks. By 24 March, a nation-wide ban on public assemblies over 20 took effect. All swimming pools, museums, libraries and bars closed, as did any businesses requiring a proximity of less than 2 metres. The impact on aggressive testing, quarantining and mild social distancing can been seen in the decrease in the effective reproduction number with $R_{\text{eff}} \geq 3.0$ on or about 13 March and decreasing to $R_{\text{eff}} \approx 0.2$, before being eliminated by 15 April. After this time there is no community transmission and R_{eff} is undefined. These effects are understated by the EpiForecast nowcast, which continues to predict an increasing value of R_{eff} from 20 April. Note that EpiForecasts does not claim their results are valid for countries such as Iceland with fewer than 100 cases per day. We include their plot here simply to illustrate that by comparison, our new process does yield sensible results for Iceland.

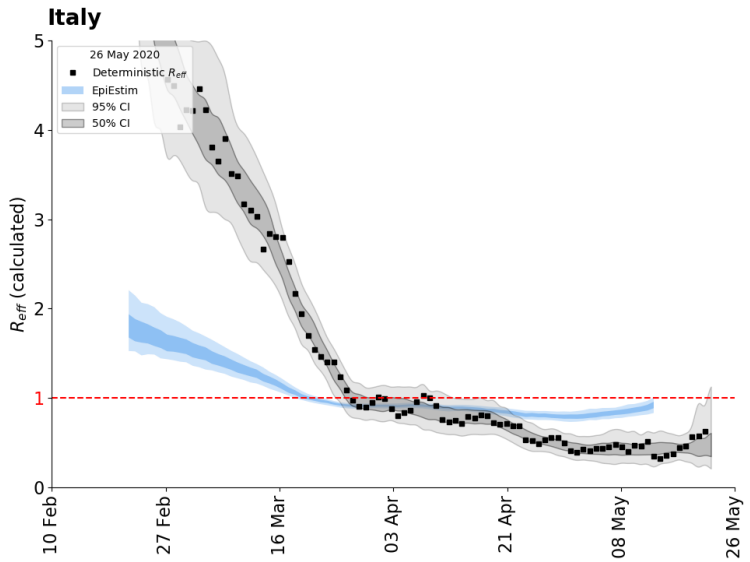

Figure 11: **Italy R_{eff} results compared with EpiForecasts.**
Italy was one of the first European states to experience COVID-19 infections, after a couple arrived from Wuhan on 23 January [16]. The Italian government response was uneven and the epidemic grew to the point of overwhelming the health system and local authorities' ability to respond, in the north of the country. Lombardy and other regions were in crisis by early March. On 10 March, prime minister Conte announced a "stay at home" order to cover all of Italy, including travel restrictions and a ban on public gatherings. On 21 March, Conte announced further restrictions within the nationwide lockdown, by halting all non-essential production, industries and businesses in Italy, following the rise in the number of new cases and deaths in the previous days. The slow and varied response of the government is reflected in the slow decrease in the effective production number, with $R_{\text{eff}} \geq 4.0$ on 28 February, decreasing to $R_{\text{eff}} \leq 1.0$ on or about 30 March. Then $R_{\text{eff}} \leq 1.0$ since that time. The EpiForecast data understates the changes in R_{eff}.

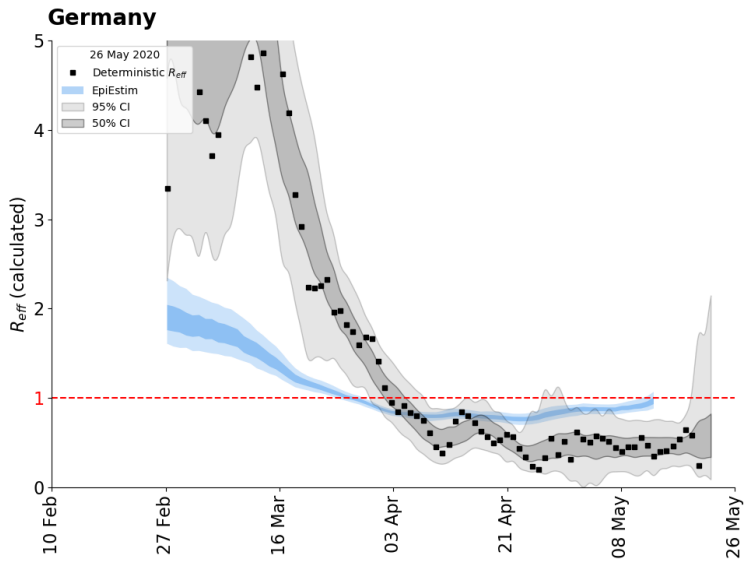

Figure 12: **Germany R_{eff} results compared with EpiForecasts.**
The first confirmed case of COVID-19 in Germany occurred on 27 January. The government response to COVID-19 was enacted at the state level, by order of federal health minister Jens Spahn on 26 February [18]. Restrictions were limited to quarantining infected cases, limited travel restrictions (notably to areas with high numbers of cases), and ad hoc cancellations of large events; however, there were no national rules. On 14 March, a new Cabinet Committee was set up and over the next week country-wide restrictions were put in place, including closing schools and religious services, shops, and businesses. Although the overall trend of the effective reproduction number R_{eff} for Germany looks similar to that of Italy, Germany achieved $R_{\text{eff}} \leq 1.0$ much more quickly. $R_{\text{eff}} \geq 4.0$ on 13 March, and it decreased to $R_{\text{eff}} \leq 1.0$ on or around 1 April, in about half the time it took Italy, demonstrating that coordinated regional responses can be as effective or more so than slow national approaches. Since that time, $R_{\text{eff}} \leq 1.0$.

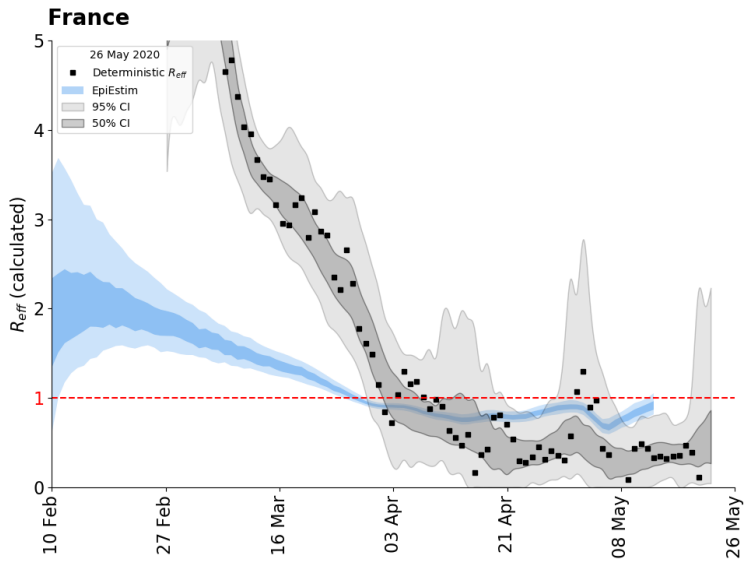

Figure 13: **France R_{eff} results compared with EpiForecasts.**
The first confirmed case of COVID-19 in Europe was on 24 January in France [19]. (Earlier cases from December last year have since been confirmed as COVID-19). After a number of clusters emerged across the country, the French government enacted restrictions on the size of gatherings, reducing the maximum allowed from 5,000 on 5 March to 100 on 14 March. Municipal elections occurred on 15 March and the official lockdown period began on 16 March, with wider restrictions on public services, institutions, and travelling. Since then, incremental increases and extensions to lockdown restrictions have been implemented. From 11 May, France began to ease lockdown restrictions in incremental steps. The incremental response of the government is reflected in the slow but steady decrease in the effective reproduction number with $R_{\text{eff}} \geq 4.0$ on 26 February, decreasing to $R_{\text{eff}} \leq 1.0$ for the first time on or about 31 March. The $R_{\text{eff}} \geq 1.0$ for two periods since 31 March between 2 and 6 April, and between 2 and 5 May. Compared to our process, the EpiForecast data understates the changes in R_{eff}.

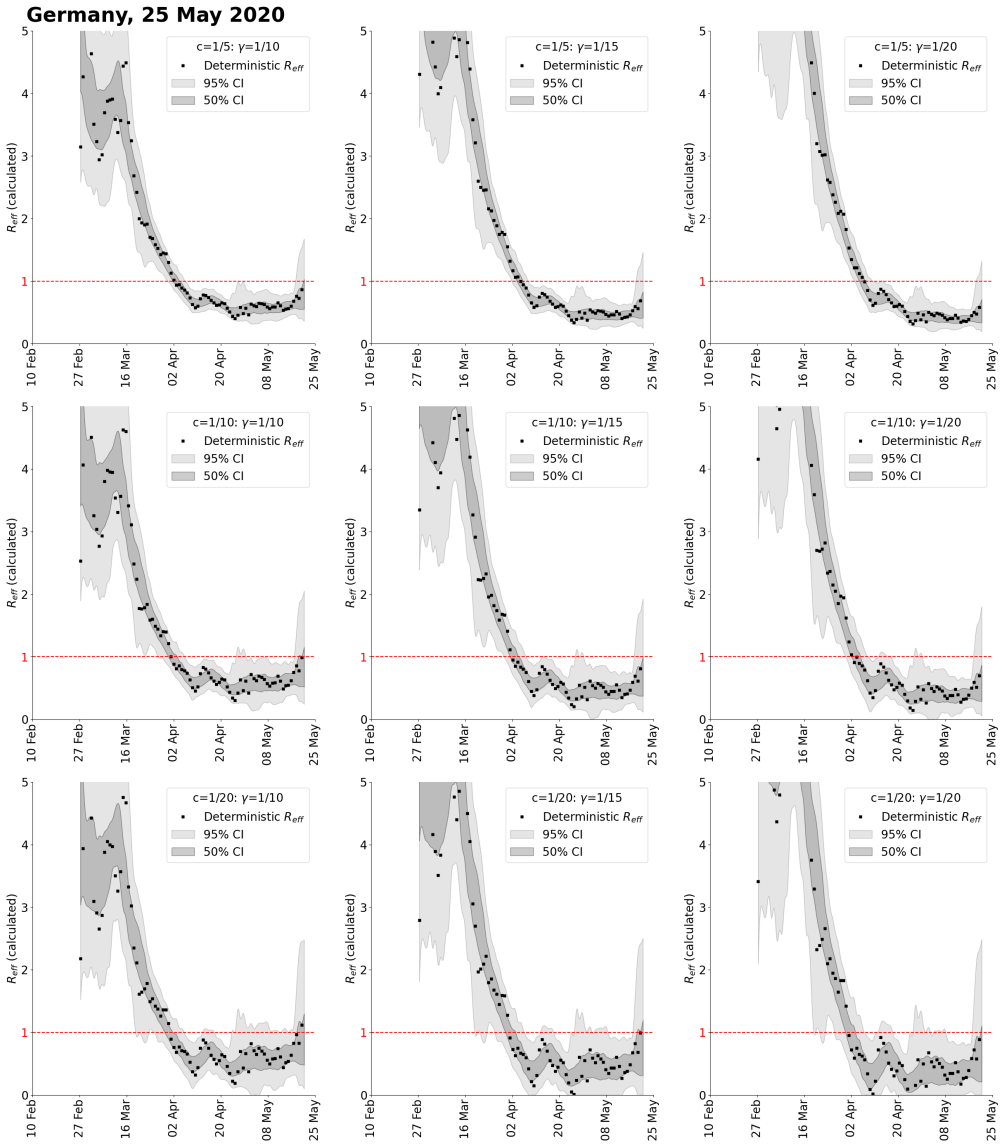

Figure 14: **Germany R_{eff} with different input parameters c and γ.**
This series of graphs displays the effect of varying the recovery parameter γ and the testing parameter c using the SIR+B method for the German daily case data. The recovery parameter γ varies through values $\{1/10, 1/15, 1/20\}$ from left to right, and c varies through $\{1/5, 1/10, 1/20\}$ from top to bottom. The calculations show that qualitatively the graphs are insensitive to the value of c and only mildly sensitive to the value of γ. Quantitatively the results are most sensitive to γ at the early stage of the contagion. The midpoint value of $\gamma = 1/15$ which we used in the earlier analyses is justified clinically on the basis that the mean time from initial infection to recovery seems to be about 15 days.

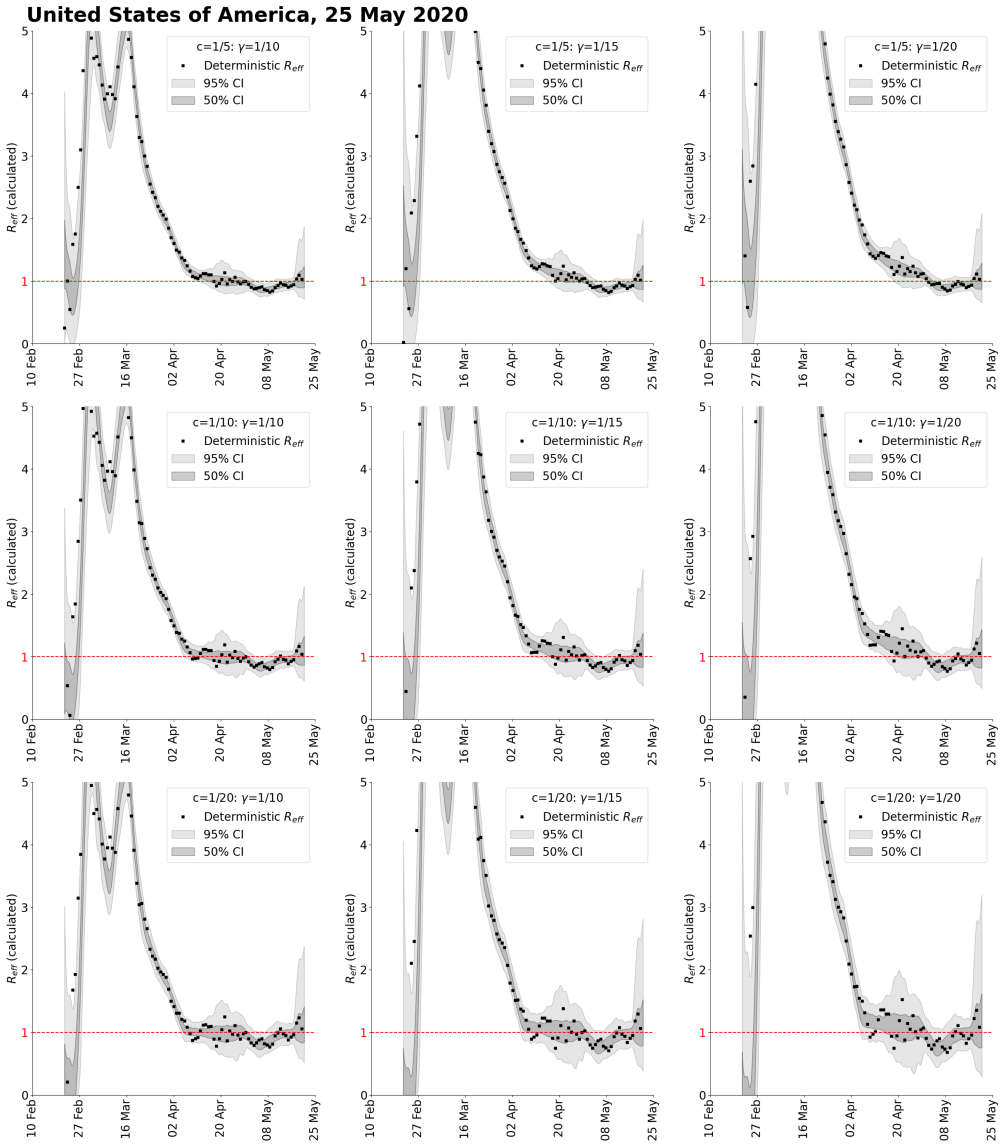

Figure 15: **United States of America R_{eff} with different input parameters c and γ.**
This series of graphs displays the effect of varying the recovery parameter γ and the testing parameter c using the SIR+B method for the United States daily case data. The recovery parameter γ varies through the values $\{1/10, 1/15, 1/20\}$ from left to right, and c varies through $\{1/5, 1/10, 1/20\}$ from top to bottom. As for the German case in Figure 14, the calculations show that qualitatively the graphs are insensitive to the value of c and only mildly sensitive to the value of γ. Quantitatively the results are most sensitive to γ at the early stage of the contagion. The midpoint value of $\gamma = 1/15$ which we used in the earlier analyses is justified clinically on the basis that the mean time from initial infection to recovery seems to be about 15 days.

A Brief overview of the method

We proceed backward through the SIR model equations (see below) to obtain a formula for the transmission coefficient multiplied by the proportion of Susceptibles in the population. From this one obtains the effective reproduction number

$$R_{\text{eff}}(t) = \frac{\beta(t)S(t)}{\gamma}.$$

This formula for $R_{\text{eff}}(t)$ can be evaluated by using the number of Infectious-Tested cases, which we obtain from reported data on the number of confirmed, tested cases per day. Sources of error and uncertainty in this deterministic stage of the process include a certain amount of guesswork in choosing the recovery parameter γ and the testing rate c, and of course errors due to noisy data.

Model failure is a possibility too. If the SIR model is not applicable, then the method for finding the reproduction number will not give sensible output. This happens when the number of new cases per day is small, which in practice we find to mean less than about 10 per day.

After the deterministic stage of the process has calculated R_{eff}, its values are fed into the second stage, which generates a large number of daily case data sets by the wild bootstrapping method. These data sets are used to generate confidence intervals for R_{eff}.

Details of the method are presented in our companion paper [5].

B SIR model equations, and parameter values

Regard the population as having fixed size N, and consider the fraction of the population falling into each of the following categories:

Susceptible, Infectious-Untested, Infectious-Tested, Removed.

This SIR system is illustrated in Figure 16.

Continuous time model

The model equations are most easily stated using continuous time, that is, for a differential equations model, and so we state them that way first. The continuous time model equations are:

$$\frac{dS}{dt} = -\beta S(I_U + I_T) \tag{1}$$

$$\frac{dI_U}{dt} = \beta S(I_U + I_T) - (c+\gamma)I_U \tag{2}$$

$$\frac{dI_T}{dt} = cI_U - \gamma I_T \tag{3}$$

$$\frac{dR}{dt} = \gamma(I_U + I_T) \tag{4}$$

As a consistency check on any simulation, one should be able to compute that

$$S + I_U + I_T + R = 1$$

for all t, since the expression on the left has derivative zero.

Notes on the model:

- The population is assumed to be homogeneous e.g. no distinctions by age.
- The model assumes random mixing among the population.
- Infectious individuals might be either symptomatic or asymptomatic.
- The transmission coefficient $\beta = \beta(t)$ is not constant. It is a function of time t.

The equations involve several parameters:

Parameter	Symbol	Value
Transmission coefficient	$\beta(t)$	to be determined
$I \to R$ transition rate	γ	(1/15) day^{-1}
Testing rate for infectious individuals	c	(1/10) day^{-1}

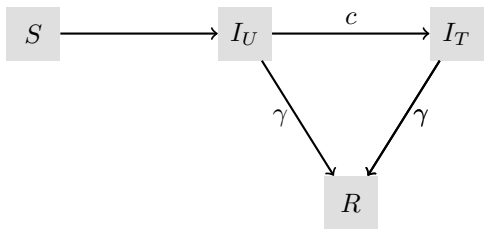

Figure 16: Progression through the SIR system: Susceptible individuals progress to Infectious-Untested, and then to either Infectious-Tested or Removed.

The value γ = (1/15) day^{-1} is chosen because the typical disease progression for COVID-19 seems to take about 15 days from initial infection. The testing rate c =(1/10) day^{-1} reflects a guess as to what fraction of Infectious-Untested individuals get tested per day. Of course, one can easily run the method with different choices for c.

Discrete time model

In the real world, new cases are reported daily, and so it makes sense to use the following discrete time analogue of the model equations, where n represents the n-th day and the time step is 1 day:

$$S(n+1) - S(n) = -\beta(n)S(n)(I_U(n) + I_T(n)) \tag{5}$$
$$I_U(n+1) - I_U(n) = \beta(n)S(n)(I_U(n) + I_T(n)) - (c+\gamma)I_U(n) \tag{6}$$
$$I_T(n+1) - I_T(n) = cI_U(n) - \gamma I_T(n) \tag{7}$$
$$R(n+1) - R(n) = \gamma(I_U(n) + I_T(n)) \tag{8}$$

As a consistency check on any simulation, one should be able to compute that

$$S + I_U + I_T + R = 1$$

for all n, since the expression on the left is independent of n (as one seems by summing the left sides of (5)–(8) and observing that the right sides sum to 0).

Effective reproduction number

The effective reproduction number is defined by

$$\boxed{R_{\text{eff}}(t) = \frac{\beta(t)S(t)}{\gamma}} \quad \text{or} \quad \boxed{R_{\text{eff}}(n) = \frac{\beta(n)S(n)}{\gamma}}, \tag{9}$$

depending whether the continuous time or discrete time model is used.

Acknowledgements

The authors acknowledge the use of open source software `NumPy` [20], `Matplotlib` [21] and `Pandas` [22].

References

[1] J. D. Murray. Mathematical Biology. I and II. An Introduction. Third edition. Interdisciplinary Applied Mathematics, 17 and 18. Springer–Verlag, New York, 2002 and 2003.

[2] A. Cori, N. M. Ferguson, C. Fraser, S. Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. Am. J. Epidemiol., 178 (2013), 1505–1512.

[3] R. N. Thompson, J. E. Stockwin, R. D. van Gaalen, J. A. Polonsky, Z. N. Kamvar, P. A. Demarsh, E. Dahlqwist, S. Li, E. Miguel, T. Jombart, J. Lessler, S. Cauchemez, A. Cori. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. Epidemics 29 (December 2019), 100356.

[4] EpiForecasts `https://epiforecasts.io/covid/`

[5] P. Grice, S. Grice, R. S. Laugesen. A new method for calculating the effective reproduction number for COVID-19. In preparation.

[6] EpiForecasts methods. https://github.com/epiforecasts/covid/blob/master/methods.Rmd

[7] European Centre for Disease Prevention and Control. Download today's data on the geographic distribution of COVID-19 cases worldwide https://www.ecdc.europa.eu/en/publications-data/download-todays-data-geographic-distribution-covid-19-cases-worldwide

[8] COVID-19 pandemic in Brazil - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Brazil

[9] COVID-19 pandemic in India - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_India

[10] COVID-19 pandemic in New Zealand - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_India

[11] Control, elimination, eradication and re-emergence of infectious diseases: getting the message right https://www.who.int/bulletin/volumes/84/2/editorial10206html/en/

[12] New Zealand's elimination strategy for the COVID-19 pandemic and what is required to make it work https://www.nzma.org.nz/journal-articles/new-zealands-elimination-strategy-for-the-covid-19-pandemic-and-what-is-required-to-make-it-work

[13] COVID-19 pandemic in the United Kingdom - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_the_United_Kingdom#Timeline

[14] COVID-19 pandemic in South Korea - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_South_Korea

[15] COVID-19 pandemic in Australia - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Australia

[16] COVID-19 pandemic in Italy - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Italy

[17] COVID-19 pandemic in Germany - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_Germany

[18] Chronology of COVID-19 pandemic in Germany - Berlin Inspector https://berlinspectator.com/2020/05/01/chronology-germany-and-the-coronavirus-2/

[19] COVID-19 pandemic in France - Wikipedia https://en.wikipedia.org/wiki/COVID-19_pandemic_in_France

[20] S. van der Walt, S. C. Colbert, G. Varoquaux. The NumPy Array: A Structure for Efficient Numerical Computation. https://ieeexplore.ieee.org/document/5725236

[21] J. D. Hunter. Matplotlib: A 2D graphics environment. https://ieeexplore.ieee.org/document/4160265/

[22] W. McKinney. Data Structures for Statistical Computing in Python. http://conference.scipy.org/proceedings/scipy2010/mckinney.html