Investigating the Impact of Asymptomatic Carriers on COVID-19 Transmission

Jacob B. Aguilar[1], Ph.D.,
Jeremy Samuel Faust[2], MD MS,
Lauren M. Westafer[3], DO, MPH, MS,
Juan B. Gutierrez[4*], Ph.D.

[1]Assistant Professor of Mathematics, Department of Mathematics and Sciences, Saint Leo University.

[2]Brigham and Women's Hospital, Department of Emergency Medicine

Instructor, Division of Health Policy and Public Health, Harvard Medical School.

[3]Assistant Professor, Assistant Professor, Department of Emergency Medicine,

University of Massachusetts Medical School-Baystate

Institute for Healthcare Delivery and Population Science.

[4]Professor of Mathematics, University of Texas at San Antonio.

*To whom correspondence should be addressed: E-mail: juan.gutierrez3@utsa.edu.

Supplemental material at: https://tinyurl.com/USA-COVID-19-Cases

Coronavirus disease 2019 (COVID-19) is a novel human respiratory disease caused by the SARS-CoV-2 virus. Asymptomatic carriers of the virus display no clinical symptoms but are known to be contagious. Recent evidence reveals that this sub-population, as well as persons with mild and undocumented disease, represent a major contributor in the propagation of this disease. However, this method of transmission frequently escapes detection by public health surveillance systems. Because of this, the currently accepted estimates of the basic reproduction number ($\mathcal{R}_0$) of the virus are inaccurate.

Further, it is unlikely that a pathogen can blanket the planet in three months with an $\mathcal{R}_0$ in the vicinity of 3, as reported (*1–6*). In this manuscript, we present a mathematical model taking into account asymptomatic and other undocumented carriers. Our results indicate that an initial value of the effective reproduction number could range from 5.5 to 26.5. Results from the first three weeks of the model exhibit exponential growth, which is in agreement with average case data collected from thirteen countries with universal health care and robust communicable disease surveillance systems; the average rate of growth in the number of reported cases is 23.3% per day during this period.

1 Background

Coronavirus disease 2019 (COVID-19) is a novel human respiratory disease caused by the SARS-CoV-2 virus. The first cases of COVID-19 disease surfaced in late December 2019 in Wuhan city, the capital of Hubei province in China. Shortly after, the virus quickly spread to several countries (*7*). On January 30, 2020 The World Health Organization (WHO) declared the virus as a public health emergency of an international scope (*8*). Twelve days later, on March 11, 2020 it was officially declared to be a global pandemic.

Asymptomatic transmission of COVID-19 has been documented (*9, 10*). The viral loads of asymptomatic carriers are similar to those in symptomatic carriers (*11*). A recent study concluded that asymptomatic and symptomatic carriers may have the same level of infectiousness (*12*). These findings demand a reassessment of the transmission dynamics of the COVID-19 outbreak that better account for asymptomatic and undocumented transmission.

The primary aim of this manuscript is to characterize the epidemiological dynamics of SARS-CoV-2 via a compartmentalized model that takes into account asymptomatic and undocumented sub-populations. The most notable result is that with the most recent data at the time of publication, SARS-CoV-2 has a large basic reproductive number $\mathcal{R}_0$ which can be estimated to fall between 5.5 and 25.4.

2 Methods

In this section we summarize the main results, and leave mathematical proofs for the appendix. Numerical estimates for the basic reproduction number follow.

2.1 Mathematical Model

The formulation of the $SEYAR$ model for the spread of SARS-CoV-2 begins with decomposing the total host population (N) into the following five epidemiological classes: susceptible human (S), exposed human (E), symptomatic human (Y), asymptomatic human (A), and recovered human (R).

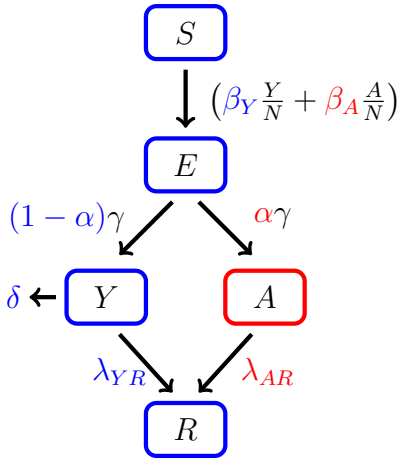


Figure 1: Schematic diagram of a COVID-19 model including an asymptomatic compartment. The arrows, except the disease-induced death (δ), represent progression from one compartment to the next. Hosts progress through each compartment subject to the rates described below.

Listed below is a $SEYAR$ dynamical system in Equation 1 describing the dynamics of COVID-19 transmission in a human population.

$$\begin{cases} \dot{S} = -\left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N}\right) S, \\ \dot{E} = \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N}\right) S - \gamma E, \\ \dot{Y} = \gamma(1-\alpha)E - (\delta + \lambda_{YR})\, Y, \\ \dot{A} = \gamma\alpha E - \lambda_{AR} A, \\ \dot{R} = \lambda_{AR} A + \lambda_{YR} Y, \end{cases} \tag{1}$$

where, $N = S + E + Y + A + R$. In a mathematical context, the reproduction number $\mathcal{R}_0$ is a threshold value that characterizes the local asymptotic stability of the underlying dynamical system at a DFE. The reproduction number arising from the dynamical system in Equation 1 is given by

$$\boxed{\mathcal{R}_0 = (1-\alpha) \cdot \beta_Y \cdot \frac{1}{\lambda_{YR} + \delta} + \alpha \cdot \beta_A \cdot \frac{1}{\lambda_{AR}}.} \tag{2}$$

As the disease-induced death rate δ is of negligible size, the reproduction number $\mathcal{R}_0$ featured in Equation 2 above admits the following natural biological interpretation:

$$\begin{aligned} \mathcal{R}_0 = &\begin{pmatrix} \text{probability of becoming} \\ \text{symptomatic upon infection} \end{pmatrix} \cdot \begin{pmatrix} \text{symptomatic} \\ \text{contact rate} \end{pmatrix} \cdot \begin{pmatrix} \text{mean symptomatic} \\ \text{infectious period} \end{pmatrix} \\ &+ \begin{pmatrix} \text{probability of becoming} \\ \text{asymptomatic upon infection} \end{pmatrix} \cdot \begin{pmatrix} \text{asymptomatic} \\ \text{contact rate} \end{pmatrix} \cdot \begin{pmatrix} \text{mean asymptomatic} \\ \text{infectious period} \end{pmatrix}. \end{aligned}$$

A mathematical proof of the calculation yielding the reproduction number $\mathcal{R}_0$ given by Equation 2 is provided in the Appendix.

The reproduction number $\mathcal{R}_0$ provides a way to measure the contagiousness of a disease. It is utilized by public health authorities to gauge the severity of an outbreak. The design and effective implementation of various intervention strategies are guided by estimates of $\mathcal{R}_0$. Established outbreaks will fade provided that interventions maintain $\mathcal{R}_0 < 1$. The reproduction number is not a biological constant corresponding to a given pathogen (*13*). In reality, the values of $\mathcal{R}_0$ fluctuate with time, and depend on environmental and socioeconomic factors.

2.2 Computation of $\mathcal{R}_0$

During the first stages of an epidemic, calculating $\mathcal{R}_0$ poses significant challenges. Evidence of this difficulty was observed in the 2009 influenza A (H1N1) virus pandemic (*14*). Particularly, the COVID-19 pandemic has a different characterization in each country in which it has spread due to differences in surveillance capabilities of public health systems, socioeconomic factors, and environmental conditions.

During the initial growth of an epidemic, Anderson et al. (*15*) derived the following formula to determine $\mathcal{R}_0$:

$$\mathcal{R}_0 = 1 + \frac{D \ln 2}{t_d}, \tag{3}$$

where D is the duration of the infectious period, and t_d is the initial doubling time. To find t_d, simply solve for t in $Y = a_0 \cdot (1+r)^t$, where $Y = 2a_0$, and $r = 23.22\%$ (the rationale for this number is explained below). Thus, $t_d = \ln 2 / \ln b$. Using the values reported on Table 1, the calculated value of the basic reproduction number using Equation 3 is $\mathcal{R}_0 = 11$, which is substantially larger than what is being reported in the literature as the COVID-19 pandemic unfolds, but should be understood as an underestimation of the true $\mathcal{R}_0$ because there is no consideration of asymptomatic carriers (*15*).

A striking characteristic of COVID-19 is the nearly perfect exponential growth reported during the first three weeks of community transmission. Figure 2 shows the number of cases reported in thirteen countries with universal health care and strong surveillance systems as of March 17, 2020. Ten of these countries are in the European zone, plus Australia, Canada and Japan. An exponential fitting for each country, conducted with the Nelder-Meade simplex algorithm (*16*), reveals an average coefficient of determination $R^2 = 0.9846 \pm 0.0164$. The average growth rate r in the exponential model $Y = a \cdot (1+r)^t$, where t is time measured in days, is $r = 23.32\%$, and the average of the initial conditions

is $a = 103$ cases. Thus, the average growth of the symptomatic compartment (Y) of COVID-19 during the first three weeks of community transmission in thirteen countries is characterized in average by the equation

$$Y = 103 \cdot 1.2332^t, \tag{4}$$

where Y_d represents the distribution of time series of reported cases, and t is time measured in days.

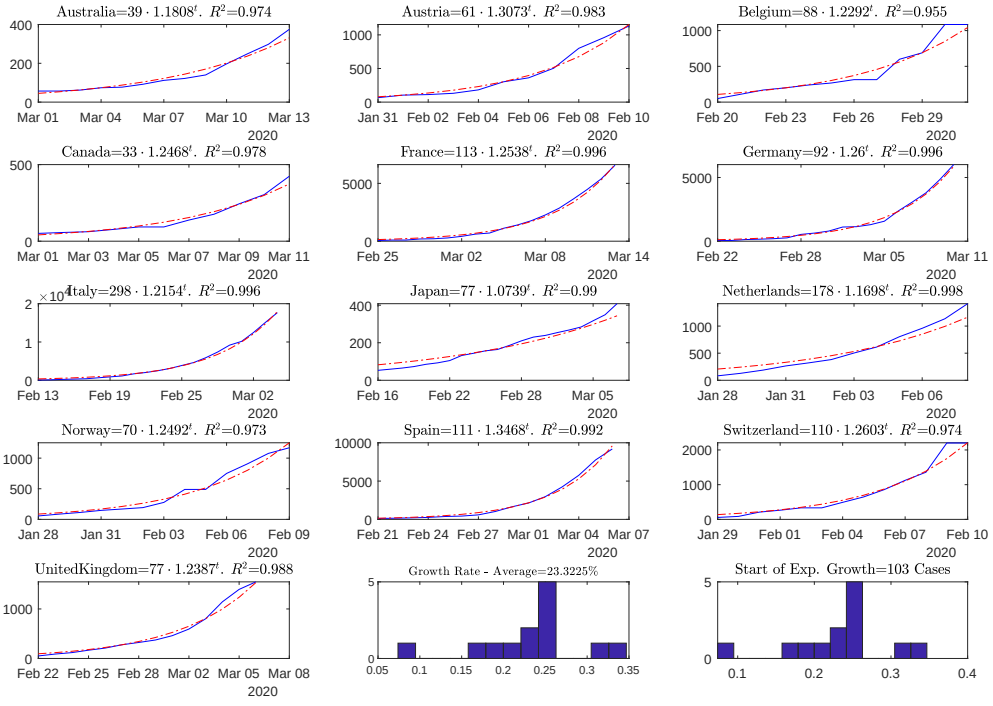


Figure 2: First three weeks (or less) of data for thirteen countries with COVID-19 cases and strong surveillance systems for communicable diseases.

There are well known challenges in attempting to fit an exponential function to epidemiological data (*17–19*). However, given the relatively slow progression of COVID-19, and the protracted incubation period, the growth of the symptomatic population can be

well characterized by an exponential function for up to three weeks.

The parameters with the greatest uncertainty at the moment of writing are λ_{YR} and λ_{AR}; hence, we calculated the range of $\mathcal{R}_0$ using the highest and lowest available values for these parameters. We did not attempt to estimate parameters based on data. To compare the output of the model to the data from the thirteen countries studied, the growth rate found in Equation 4 was superimposed on the model. The initial condition a_0 in the exponential function $Y = a_0 \cdot (1 + r)^t$ was fitted tot he dynamical system with the Nelder-Meade simplex algorithm (*16*). It is important to emphasize that fitting the initial value simple creates a translation of the curve. It is, therefore, remarkable that the function that describes the average behavior of the first three weeks around the world, presents a nearly perfect fit to the dynamical system using parameters that were measured in multiple settings by different groups.

Table 1: Model Parameters

Parameter	**Description**	**Dimension**	**Median(95% CIs) Value or Range**	**Source**
β_Y	effective contact rate from symptomatic to susceptible	$days^{-1}$	$1.12(1.07, 1.18)$	(*20*)
ϕ	relative transmission probability	n/a	$0.55(0.49, 0.63)$	(*20*)
β_A	effective contact rate from asymptomatic to susceptible	$days^{-1}$	$\phi\beta_Y$	(*20*)
γ^{-1}	mean serial period	days	$5.1(4.5, 5.8)$	(*21*)
α	probability of becoming asymptomatic upon infection	n/a	$0.86(0.82, 0.9)$	(*20, 22*)
λ_{YR}^{-1}	mean symptomatic infectious period	*days*	$[8, 37]$	(*23*)
λ_{AR}^{-1}	mean asymptomatic infectious period	*days*	$[8, 37]$	Assumed
δ	disease-induced death rate	$days^{-1}$	$0.032(1 - \alpha)$	(*24*)

At the time of writing, there are not reliable estimations for the infectious period of asymptomatic carriers. In absence of data, we decided to use the estimates of infec-

tiousness measured for symptomatic patients. All other parameters have been directly measured, or derived from measures.

3 Results

Figure 3 shows a calculation of the SEYAR model using the parameters reported in Table 1. This representation of the progression of the disease must be understood as a theoretical development; in reality, the progression of an epidemic depends on a multitude of factors that necessarily result in deviations from this ideal case.

Changes in behavioral patterns in response to an outbreak have an effect on the advance of the given pathogen. As people gain awareness of the presence of an infectious disease near their communities, a portion will take measures in order to reduce their susceptibility. An example of this behavior corresponding to the COVID-19 pandemic is that of social distancing. Indeed, the cancellation of events likely to attract crowds, the closing of schools, and working from home will have a drastic impact on the size of the susceptible population at a given time. Figure 3. shows time series corresponding to the upper and lower estimations of the basic reproduction number, along with intervention simulations for each scenario. Figure 4 shows the variation of $\mathcal{R}_0$ with respect to the symptomatic and asymptomatic mean infectious periods, λ_{YR}^{-1} and λ_{AR}^{-1}.

The size of the COVID-19 reproduction number documented in the literature is relatively small. Our estimates indicate that $\mathcal{R}_0$is likely to be in the interval from 5.5 to 25.4, in the case that the undocumented infection and asymptomatic infection sub-population is accounted for. Considering two scenarios, one with the upper estimation of $\mathcal{R}_0$ and one with the lower, the peak of symptomatic infections is reached in around 31 days, with percentages of the entire population being symptomatic at 9% and 4%, respectively. These scenarios are shown in Figure 3 below. The effects of interventions are shown for

illustration purposes only. There is no data to calibrate an intervention scenario at the time of writing.

4 Discussion

The calculation of $\mathcal{R}_0$ poses significant challenges during the first stages of the COVID-19 epidemic. This is due to paucity and timing of surveillance data, different methodological approaches to data collection, and different guidelines for testing. Estimates vary greatly: 0.3 (*25*), 2.28 (*26*), 2.38 (*27*), 3.28 (*28*), and others. However, none of these studies takes into consideration the possibility of asymptomatic carriers.

The time series of symptomatic individuals provided by the SEYAR model can inform the likely progression of the disease. The compartment Y must be considered as an upper bound for the progression of the COVID-19 pandemic, that is, what surveillance systems could observe in absence of public health interventions and behavior modification. However, as the COVID-19 pandemic evolves, governments around the world are taking drastic steps to limit community spread. This will necessarily dampen the growth of the disease. The SEYAR model captured faithfully the first stages of the pandemic, and remains a stark reminder of what the cost of inaction could be. It can be used as a tool to explore multiple scenarios corresponding to different interventions.

The parameters with the highest degree of uncertainty are the mean infectious periods for symptomatic (λ_{YR}^{-1}) and asymptomatic (λ_{AR}^{-1}) subjects. If we consider the median of the other parameters to be correct (there is more data), then the mean infectious period of a symptomatic patient should be 4.9 days, and the mean infectious period of an asymptomatic should be 4.1 days. These numbers are not consistent with what is being observed on the ground at the time of writing. If we reduced α, the probability of becoming asymptomatic upon infection, to less than 0.86, e.g. $\alpha = 0.5$, then the mean

infectious period of a symptomatic patient would be 3.7 days, and the mean infectious period of an asymptomatic carrier would be 3.1 days. In reality, patients are infectious before the onset of symptoms, and the disease lasts longer than 3 days in symptomatic patients. The necessary conclusion is that via a computational *reductio ad absurdum*, in tandem with the information we have today, $\mathcal{R}_0$ cannot be near 3.

5 Conclusion

It is unlikely that a pathogen that blankets the planet in three months can have a basic reproduction number in the vicinity of 3, as it has been reported in the literature (*1*–*6*). In juxtaposition to the SARS-CoV epidemic of 2003 (*29*), where only symptomatic individuals were capable of transmitting the disease, asymptomatic carriers of the SARS-CoV-2 virus are capable of transmission to the same degree as symptomatic (*11*). In a public health context, the silent threat posed by the presence of asymptomatic and other undocumented carriers in the population results in the COVID-19 pandemic being much more difficult to control. SARS-CoV-2 is probably among the most contagious pathogens known, a phenomenon most likely driven by the infectiousness of the asymptomatic sub-population.

The value of $\mathcal{R}_0$ has to be understood as threshold parameter that helps understand whether a disease will spread, and it varies substantially per locality depending on how public health officials have communicated the risk to the general public, general beliefs and (dis)information available to the population, and other socioeconomic and environmental factors. Our goal with this article was not to conduct a rigorous statistical analysis, which is a moot point so early in the growth of the COVID-19 pandemic. Instead, we can apply this mean field estimate to different locations to provide an estimation of the potential impact of the disease.

This study shows that the population of individuals with asymptomatic and undocumented COVID-19 infections are driving the growth of the pandemic. The value of $\mathcal{R}_0$ we calculated is at least double and up to one order of magnitude larger than the estimates that have been communicated in the literature up to this point.

A Appendix

The simplified dynamical system described in Equation 1 was computed with the values listed on Table 1. A literature search was conducted to determine values for these parameters. The definitions used for the parameters are as follows:

- Effective contact rate from symptomatic to susceptible sub-population (β_Y): The transmission rate corresponding to symptomatic infections accounting for effective contacts per unit time between symptomatic and susceptible individuals.

- Relative transmission probability (ϕ): A transmission reduction factor corresponding to asymptomatic infections.

- Effective contact rate from asymptomatic to susceptible sub-population (β_A): The transmission rate corresponding to asymptomatic infections accounting for effective contacts per unit time between asymptomatic and susceptible individuals.

- Mean serial period (γ^{-1}): Mean number of days between exposure to COVID-19 source and development of transmissibility. Since the serial period is shorter than the incubation period, hosts are capable of pre-symptomatic transmission. This parameter is usually called the mean incubation period in the literature.

- Probability of becoming asymptomatic upon infection (α): Transmission factor accounting for the asymptomatic sub-population.

- Mean symptomatic infectious period (λ_{YR}^{-1}): Mean number of days a patient who developed COVID-19 symptoms had viral shedding. For both the mean asymptomatic and symptomatic infectious periods we assumed viral shedding was synonymous with transmissibility.

- Mean asymptomatic infectious period (λ_{AR}^{-1}): Mean number of days a patient who never developed symptoms had viral shedding.

- Disease-induced death rate (δ): The rate of fatality caused by disease. It is taken to be the quotient of disease-induced deaths and confirmed cases. Since there are unreported and asymptomatic cases, δ is multiplied by $1-\alpha$.

Listed below is the generalized $SEYAR$ dynamical system in Equation 5 which falls into the class of models covered by Aguilar and Gutierrez (2020) (*30*), see Figure 5.

$$\begin{cases} \dot{S} = \Lambda + \lambda_{RS} R - \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N} + \xi\right) S, \\ \dot{E} = \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N}\right) S - (\gamma + \xi) E, \\ \dot{Y} = \gamma(1-\alpha) E - (\xi + \delta + \lambda_{YR}) Y + \lambda_{AY} A, \\ \dot{A} = \gamma \alpha E - (\lambda_{AR} + \lambda_{AY} + \xi) A, \\ \dot{R} = \lambda_{AR} A + \lambda_{YR} Y - (\lambda_{RS} + \xi) R, \end{cases} \tag{5}$$

where, $N = S + E + Y + A + R$. The demographic parameters Λ and ξ denote the human recruitment and mortality rates, respectively. While λ_{AY} and λ_{RS} are the asymptomatic to symptomatic transition and relapse rates, respectively.

It is worth mentioning that for a basic $SEIR$ model, where there is only one infected compartment, the progression rate from the susceptible to the exposed class λ_{SE} is equal to the product of the effective contact rate β and the proportion of infected individuals $\frac{I}{N}$, so that

$$\lambda_{SE} = \beta \frac{I}{N}.$$

In our model, we decompose the infected compartment into symptomatic and asymptomatic sub-compartments. Due to this decomposition, the progression rate is given by the weighted sum

$$\lambda_{SE} = \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N} \right).$$

Disease-Free Equilibrium (DFE) points are solutions of a dynamical system corresponding to the case where no disease is present in the population.

Lemma 1. *(Reproduction Number for the SEYAR COVID-19 Model). Define the following quantity*

$$\mathcal{R}_0 := \frac{\gamma}{\gamma + \xi} \left(\frac{\beta_Y}{\delta + \lambda_{YR} + \xi} \left(\frac{\alpha \lambda_{AY}}{\lambda_{AR} + \lambda_{AY} + \xi} - (\alpha - 1) \right) + \frac{\alpha \beta_A}{\lambda_{AR} + \lambda_{AY} + \xi} \right). \quad (6)$$

Then, the DFE $\mathbf{w}^\star$ *for the SEYAR model in Equation 5 is locally asymptotically stable provided that* $\mathcal{R}_0 < 1$ *and unstable if* $\mathcal{R}_0 > 1$.

Proof. We order the compartments so that the first four correspond to the infected subpopulations and denote $\mathbf{w} = (E, Y, A, R, S)^T$. The corresponding DFE is

$$\mathbf{w}^\star = \left(0, 0, 0, 0, \frac{\Lambda}{\xi} \right)^T.$$

Utilizing the next generation method developed by Van den Driessche and Watmough (*31*), system in Equation 5 is rewritten in the following form

$$\dot{\mathbf{w}} = \Phi(\mathbf{w}) = \mathcal{F}(\mathbf{w}) - \mathcal{V}(\mathbf{w}),$$

where $\mathcal{F} := (\mathcal{F}_1, \ldots, \mathcal{F}_5)^T$ and $\mathcal{V} := (\mathcal{V}_1, \ldots, \mathcal{V}_5)^T$, or more explicitly

$$\begin{pmatrix} \dot{E} \\ \dot{Y} \\ \dot{A} \\ \dot{R} \\ \dot{S} \end{pmatrix} = \begin{pmatrix} \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N}\right) S \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} - \begin{pmatrix} (\gamma+\xi)E \\ -\gamma\left(1-\alpha\right)E + \left(\xi+\delta+\lambda_{YR}\right)Y - \lambda_{AY}A \\ -\gamma\alpha E + \left(\lambda_{AR}+\lambda_{AY}+\xi\right)A \\ -\lambda_{AR}A - \lambda_{YR}Y + \left(\lambda_{RS}+\xi\right)R \\ -\Lambda - \lambda_{RS}R + \left(\beta_Y \frac{Y}{N} + \beta_A \frac{A}{N} + \xi\right)S \end{pmatrix}.$$

The matrix $\mathcal{V}$ admits the decomposition $\mathcal{V} = \mathcal{V}^- - \mathcal{V}^+$, where the component-wise definition is inherited. In a biological context, $\mathcal{F}_i$ is the rate of appearance of new infections in compartment i, $\mathcal{V}_i^+$ stands for the rate of transfer of individuals into compartment i by any other means and $\mathcal{V}_i^-$ is the rate of transfer of individuals out of compartment i. Now, let F and V be the following sub-matrices of the Jacobian of the above system, evaluated at the solution $\mathbf{w}^\star$

$$F = \left(\frac{\partial \mathcal{F}_i}{\partial x_j}\Big|_{\mathbf{w}^\star}\right)_{1\le i,j\le 3} = \begin{pmatrix} 0 & \beta_Y & \beta_A \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

and

$$V = \left(\frac{\partial \mathcal{V}_i}{\partial x_j}\Big|_{\mathbf{w}^\star}\right)_{1\le i,j\le 3} = \begin{pmatrix} (\gamma+\xi) & 0 & 0 \\ \gamma\left(\alpha-1\right) & \left(\xi+\delta+\lambda_{YR}\right) & -\lambda_{AY} \\ -\gamma\alpha & 0 & \left(\lambda_{AR}+\lambda_{AY}+\xi\right) \end{pmatrix}.$$

A direct calculation shows that

$$V^{-1} = \begin{pmatrix} (\gamma+\xi)^{-1} & 0 & 0 \\ -\frac{\gamma((\alpha-1)(\lambda_{AR}+\xi)-\lambda_{AY})}{(\gamma+\xi)(\xi+\delta+\lambda_{YR})(\xi+\lambda_{AY}+\lambda_{AR})} & \left(\xi+\delta+\lambda_{YR}\right)^{-1} & \lambda_{AY}\left(\left(\xi+\delta+\lambda_{YR}\right)\left(\xi+\lambda_{AY}+\lambda_{AR}\right)\right)^{-1} \\ \gamma\alpha\left(\left(\gamma+\xi\right)\left(\lambda_{AR}+\lambda_{AY}+\xi\right)\right)^{-1} & 0 & \left(\lambda_{AR}+\lambda_{AY}+\xi\right)^{-1} \end{pmatrix}$$

and FV^{-1} is given by the following matrix

$$\begin{pmatrix} \frac{\gamma}{(\gamma+\xi)(\lambda_{AR}+\lambda_{AY}+\xi)}\left(-\frac{\beta_Y((\alpha-1)(\lambda_{AR}+\xi)-\lambda_{AY})}{\delta+\lambda_{YR}+\xi} + \beta_A\alpha\right) & \beta_Y\left(\delta+\lambda_{YR}+\xi\right)^{-1} & \frac{1}{\lambda_{AR}+\lambda_{AY}+\xi}\left(\frac{\beta_Y\lambda_{AY}}{\delta+\lambda_{YR}+\xi} + \beta_A\right) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Let $\mathcal{I}$ denote the 3×3 identity matrix, so that the characteristic polynomial $P(\lambda)$ of the matrix FV^{-1} is given by

$$\begin{aligned} P(\lambda) &= \det\left(FV^{-1} - \lambda\mathcal{I}\right), \\ &= \lambda^2\left(\lambda - \left(\frac{\gamma\beta_Y}{(\gamma+\xi)(\delta+\lambda_{YR}+\xi)}\left(\frac{\alpha\lambda_{AY}}{\lambda_{AR}+\lambda_{AY}+\xi}+1-\alpha\right)+\frac{\gamma\alpha\beta_A}{(\gamma+\xi)(\lambda_{AR}+\lambda_{AY}+\xi)}\right)\right). \end{aligned}$$

The solution set $\{\lambda_i\}_{1\leq i\leq 3}$ is given by

$$\left\{0, 0, \frac{\gamma\beta_Y}{(\gamma+\xi)(\delta+\lambda_{YR}+\xi)}\left(\frac{\alpha\lambda_{AY}}{\lambda_{AR}+\lambda_{AY}+\xi}+1-\alpha\right)+\frac{\gamma\alpha\beta_A}{(\gamma+\xi)(\lambda_{AR}+\lambda_{AY}+\xi)}\right\}.$$

Therefore, the reproduction number for the *SEYAR* model in Equation 5 is given by

$$\begin{aligned} \mathcal{R}_0 &:= \rho\left(FV^{-1}\right), \\ &= \max_{1\leq i\leq 3}\{\lambda_i\}, \\ &= \frac{\gamma\beta_Y}{(\gamma+\xi)(\delta+\lambda_{YR}+\xi)}\left(\frac{\alpha\lambda_{AY}}{\lambda_{AR}+\lambda_{AY}+\xi}+1-\alpha\right)+\frac{\gamma\alpha\beta_A}{(\gamma+\xi)(\lambda_{AR}+\lambda_{AY}+\xi)}, \\ &= \frac{\gamma}{\gamma+\xi}\left(\frac{\beta_Y}{\delta+\lambda_{YR}+\xi}\left(\frac{\alpha\lambda_{AY}}{\lambda_{AR}+\lambda_{AY}+\xi}-(\alpha-1)\right)+\frac{\alpha\beta_A}{\lambda_{AR}+\lambda_{AY}+\xi}\right). \end{aligned}$$

The proof of the lemma regarding the local asymptotic stability of the DFE $\mathbf{w}^\star$ corresponding to the *SEYAR* model in Equation 5 is now complete after invoking Theorem 2 reported by Van den Driessche and Watmough (2002) (*31*). □

The expression in Equation 2 corresponds to the absence of the demographic parameter and asymptomatic to symptomatic transition rate, i.e. $\xi = \lambda_{AY} = 0$. This specific case corresponds to the DFE solution given by $\mathbf{v}^\star = (0, 0, 0, 0, N(0))^T$. A verification of the calculation yielding the reproduction number $\mathcal{R}_0$ given by Equation 6 is provided in the electronic supplementary material.

The reproduction number $\mathcal{R}_0$ shown in Equation 2 arising from our model admits a natural biological interpretation. To guide this discussion, it is pertinent to refer to the original epidemic model proposed by W. O. Kermack and A. G. McKendrick in 1927 (*32*), see Figure 6 below, has the corresponding dynamical system

$$\begin{cases} \dot{S} = -\beta \frac{I}{N} S, \\ \dot{I} = \beta \frac{I}{N} S - \omega I, \\ \dot{R} = \omega I. \end{cases} \tag{7}$$

Epidemiologically speaking, the basic reproduction number is the average number of secondary infections generated by a single infection in a completely susceptible population. It is proportional to the product of infection/contact (a), contact/time (b) and time/infection (c). The quantity a is the infection probability between susceptible and infectious individuals, b is the mean contact rate between susceptible and infectious individuals and c is the mean duration of the infectious period.

The case of an increasing infected sub-population corresponds to the occurrence of an epidemic. This happens provided that $\dot{I} = \beta \frac{I}{N} S - \omega I > 0$ or $\frac{\beta}{\omega} \frac{S}{N} > 1$. Under the assumption that in the beginning of an epidemic, virtually the total population is susceptible, that is $\frac{S}{N} \approx 1$. As a result, we arrive at the following equivalent condition

$$\mathcal{R}_0 := \frac{\beta}{\omega} > 1.$$

The parameter β in Figure 6 is equal to ab and ω is equal to c^{-1}. This combination of parameters stands to reason as it is a ratio of the effective contact rate β and the mean infectious period ω^{-1}.

Since the disease-induced death rate $\delta \approx 0$, the reproduction number in Equation 2 for our model has a similar natural interpretation as the sum of ratios consisting of the effective contact rates β_Y, β_A and mean infectious periods λ_{YR}^{-1}, λ_{AR}^{-1} for the symptomatic and asymptomatic sub-populations, weighted with the probabilities of becoming symptomatic

$(1-\alpha)$ or asymptomatic α upon infection.

The effective reproduction number $\mathcal{R}_0(t)$ takes into consideration the susceptibility of the population,

$$\mathcal{R}_0(t) := \frac{\mathcal{R}_0}{N(t)} S(t). \tag{8}$$

It is defined to be the average number of secondary cases generated by a typical case. A decrease in the susceptible population overtime will cause a corresponding decrease in the values of the reproduction number. It directly follows by Equation 8 that $\mathcal{R}_0(0) = \mathcal{R}_0$, as initially the total human population is assumed to be susceptible. The plot of $\mathcal{R}_0(t)$ is similar to the plot of the susceptible portion, featured in Figure 3. This is reasonable since Equation 8 implies that $\mathcal{R}_0(t)$ is proportional to $S(t)$. Since $\delta \approx 0$, the total population $N(t)$ varies little within a tight envelope around the initial susceptible population $S(0)$. This is easily observable upon inspection of the dynamical system in Equation 1, as it is clear that

$$N(t) = S(0) - \delta \int_0^t Y(\zeta) d\zeta.$$

References

1. J. M. Read, J. R. Bridgen, D. A. Cummings, A. Ho, C. P. Jewell, *medRxiv* (2020).

2. T. Liu, *et al.* (2020).

3. M. Majumder, K. D. Mandl, *China (January 23, 2020)* (2020).

4. Z. Cao, *et al.*, *medRxiv* (2020).

5. S. Zhao, *et al.*, *International Journal of Infectious Diseases* **92**, 214 (2020).

6. N. Imai, *et al.*, *Reference Source* (2020).

7. W.-j. Guan, *et al.*, *New England Journal of Medicine* (2020).

8. World Health Organization, Coronavirus disease (covid-19) outbreak.

9. K. Mizumoto, K. Kagaya, A. Zarebski, G. Chowell, *medRxiv* (2020).

10. Z. Hu, *et al.*, *Science China Life Sciences* pp. 1–6 (2020).

11. L. Zou, *et al.*, *New England Journal of Medicine* (2020).

12. C. R. MacIntyre, *Global Biosecurity* **1** (2020).

13. P. L. Delamater, E. J. Street, T. F. Leslie, Y. T. Yang, K. H. Jacobsen, *Emerging infectious diseases* **25**, 1 (2019).

14. H. Nishiura, G. Chowell, M. Safan, C. Castillo-Chavez, *Theoretical Biology and Medical Modelling* **7**, 1 (2010).

15. R. Anderson, G. Medley, R. May, A. Johnson, *Mathematical Medicine and Biology: a Journal of the IMA* **3**, 229 (1986).

16. J. C. Lagarias, J. A. Reeds, M. H. Wright, P. E. Wright, *SIAM Journal on optimization* **9**, 112 (1998).

17. G. Chowell, C. Viboud, *Infectious disease modelling* **1**, 71 (2016).

18. G. Chowell, L. Sattenspiel, S. Bansal, C. Viboud, *Physics of life reviews* **18**, 114 (2016).

19. G. Chowell, C. Viboud, L. Simonsen, S. M. Moghadas, *Journal of The Royal Society Interface* **13**, 20160659 (2016).

20. R. Li, *et al.*, *Science* (2020).

21. S. A. Lauer, *et al.*, *Annals of Internal Medicine* (2020).

22. H. Nishiura, *et al.*, The rate of underascertainment of novel coronavirus (2019-ncov) infection: Estimation using japanese passengers data on evacuation flights (2020).

23. F. Zhou, *et al.* (2020).

24. D. Baud, *et al.*, *The Lancet Infectious Diseases* (2020).

25. P. Wu, *et al.*, *Eurosurveillance* **25** (2020).

26. S. Zhang, *et al.*, *International Journal of Infectious Diseases* (2020).

27. R. Li, *et al.*, *medRxiv* (2020).

28. Y. Liu, A. A. Gayle, A. Wilder-Smith, J. Rocklöv, *Journal of travel medicine* (2020).

29. R. M. Anderson, *et al.*, *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences* **359**, 1091 (2004).

30. J. B. Aguilar, J. B. Gutierrez, *Bulletin of Mathematical Biology* **82**, 42 (2020).

31. P. Van den Driessche, J. Watmough, *Mathematical biosciences* **180**, 29 (2002).

32. W. Kermack, A. Mckendrick, *Proc Roy Soc London* **A 115**, 700 (1927).

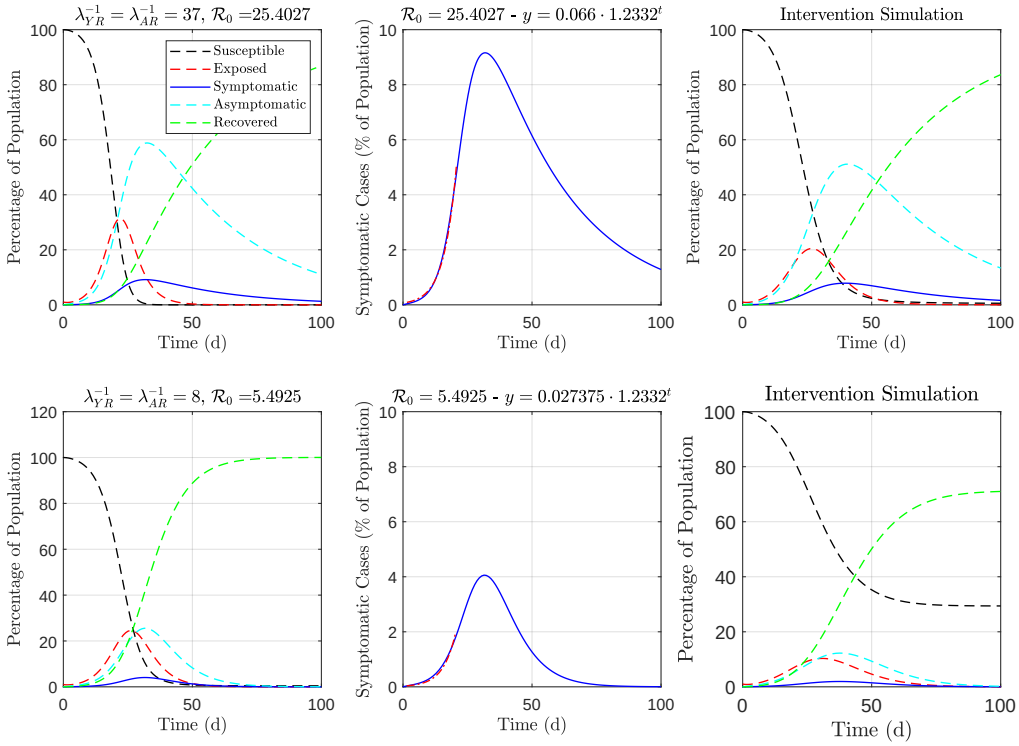


Figure 3: Numerical implementation of a SEYAR model with the parameters listed on Table 1. The left-most panels show the time series corresponding to upper and lower estimations of the COVID-19 basic reproduction number $\mathcal{R}_0$. The center panels are times series of the symptomatic compartment, with the exponential function (red dashed line) whose parameters are the average of the thirteen countries studied superimposed on the first three weeks of the outbreak. The right-most panels show intervention simulations of equal intensity computed limiting the contact between the susceptible and infected populations; at the time of writing there is not data available to calibrate an intervention model.

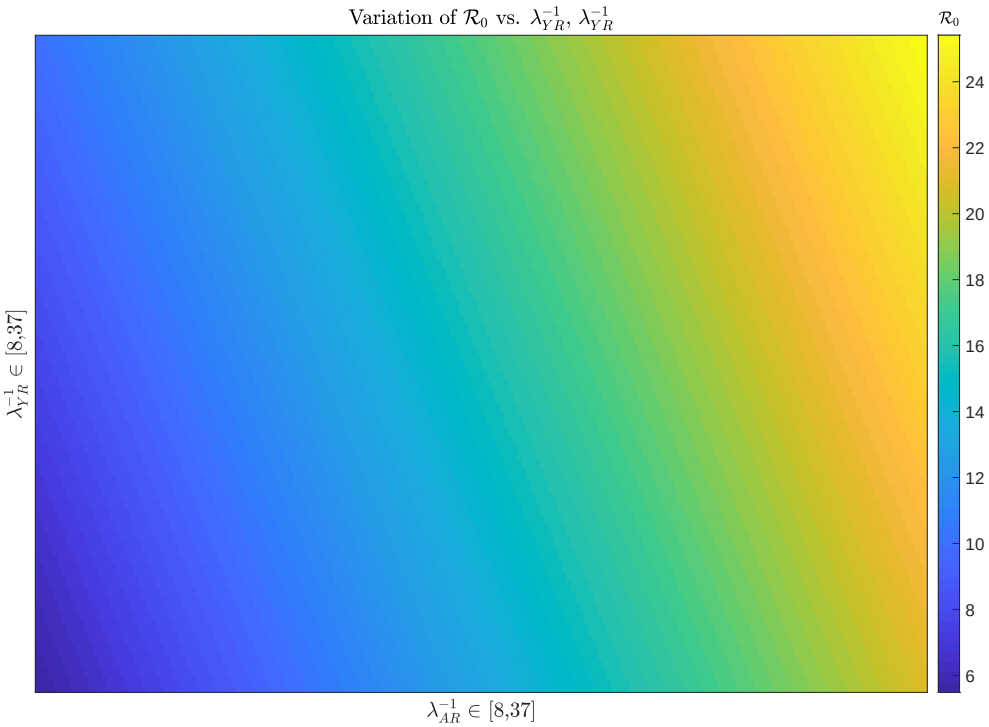


Figure 4: Heat map showing the variation of the basic reproduction number $\mathcal{R}_0$ with respect to the asymptomatic and symptomatic infectious periods λ_{AR}^{-1} and λ_{YR}^{-1}, respectively.

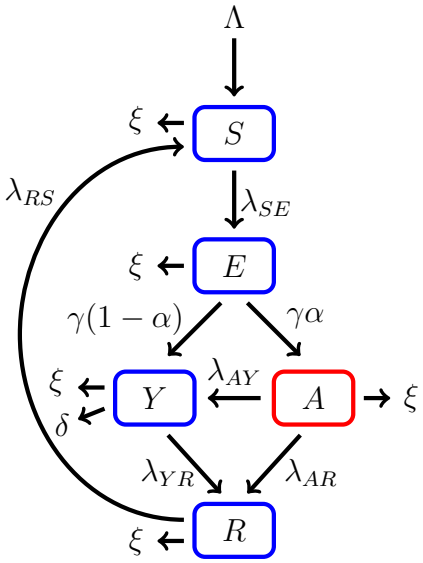


Figure 5: This figure is a schematic diagram of a generalized COVID-19 model including an asymptomatic compartment. The longer arrows represent progression from one compartment to the next. Hosts enter the susceptible compartment either through birth of migration and then progress through each additional compartment subject to the rates described above.

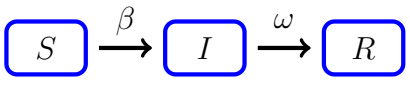
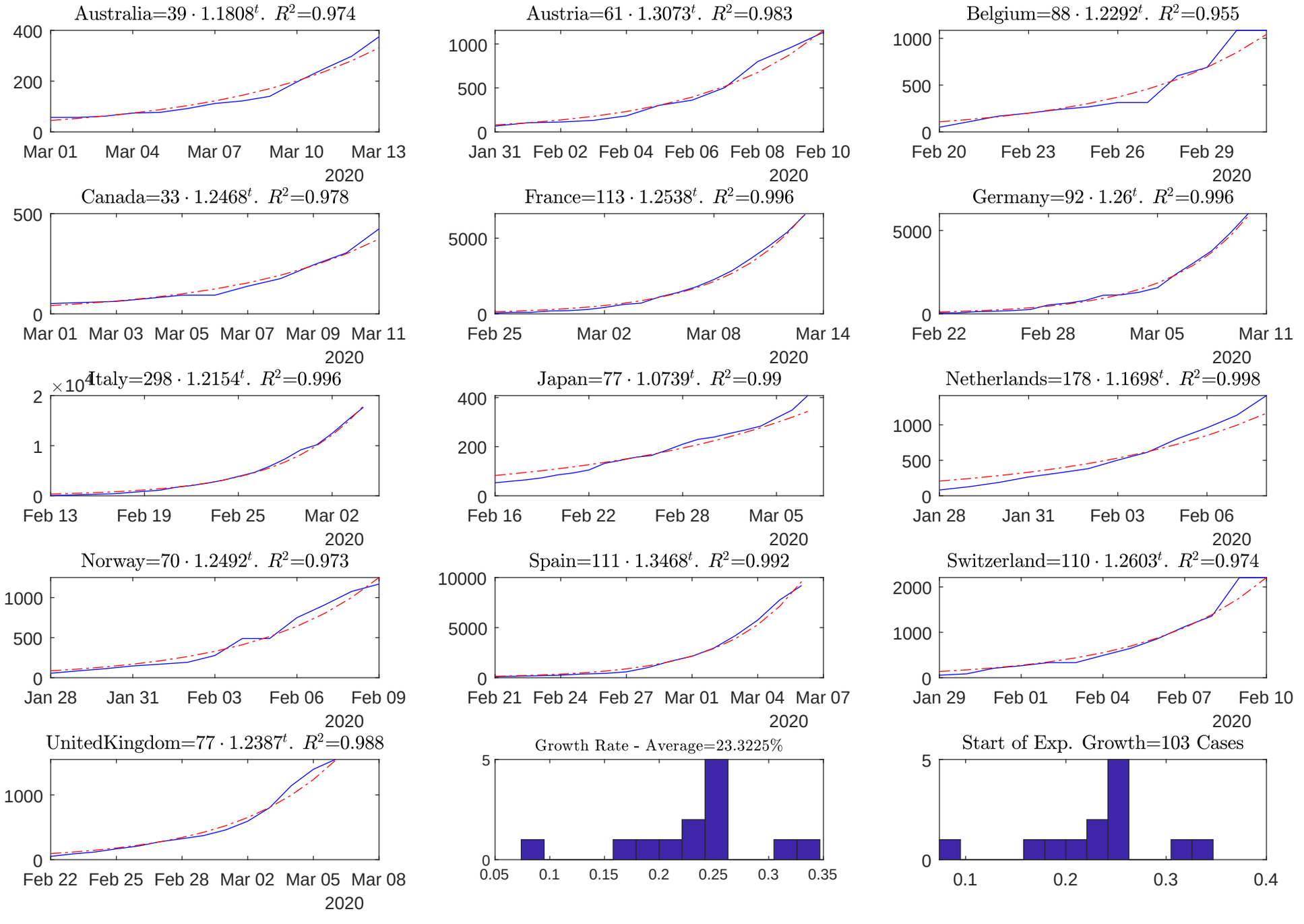
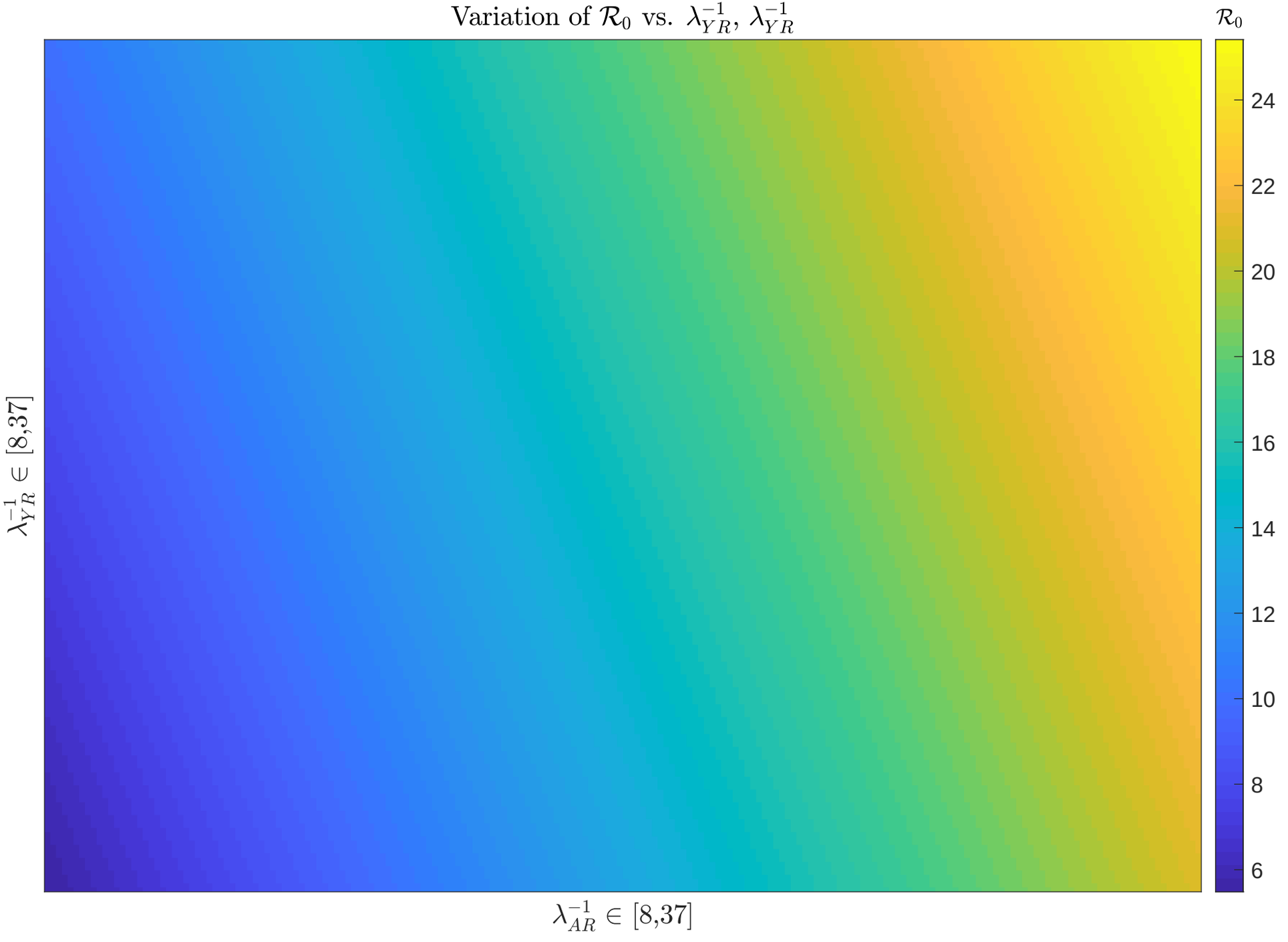
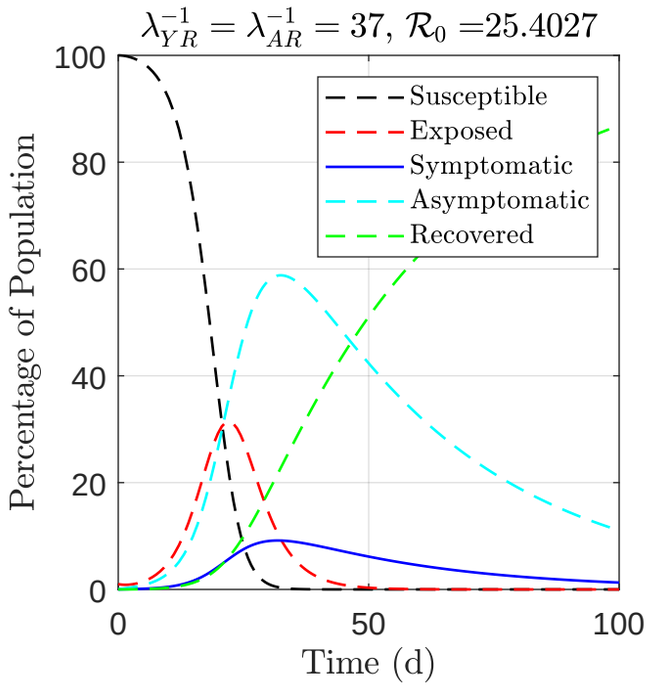
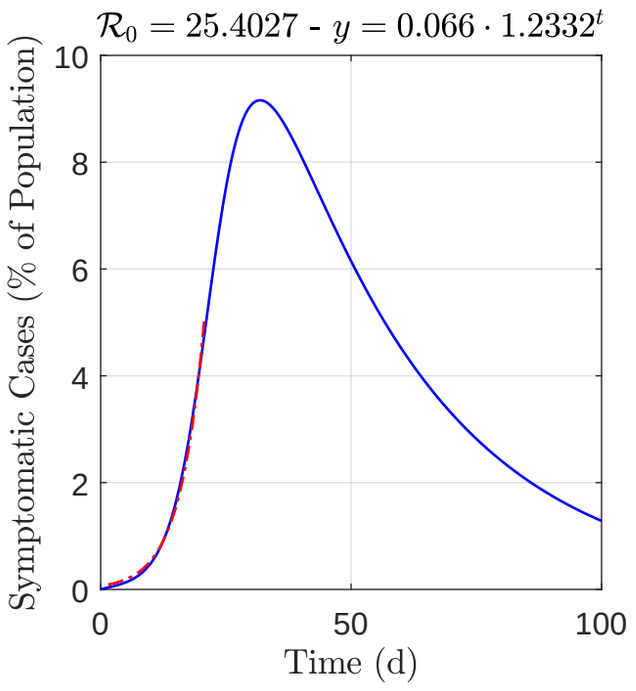


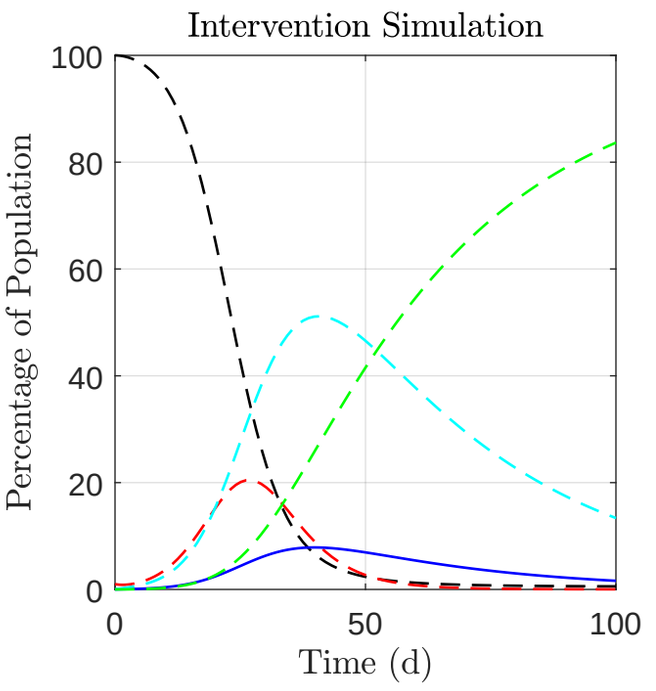
Figure 6: This figure is a schematic diagram of a SIR model consisted of three compartments, namely: susceptible (S), infected (I) and recovered (R). Humans progress through each compartment subject to the rates described above.

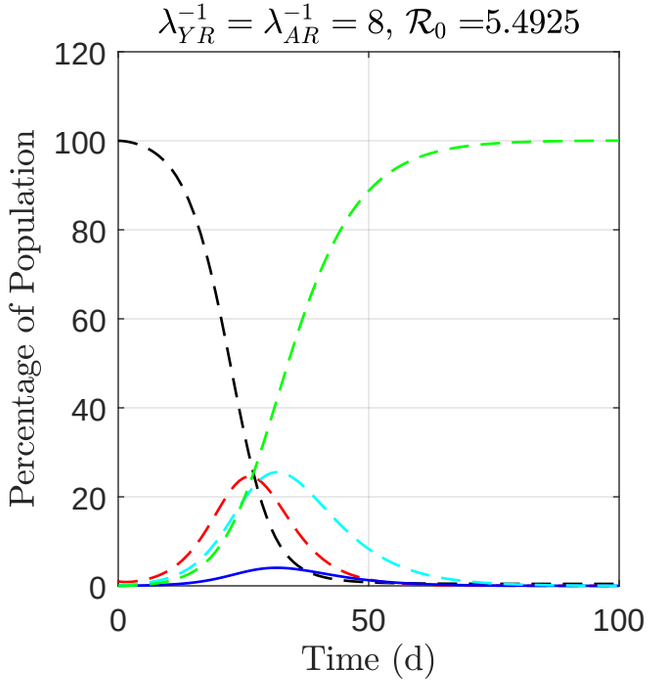

Australia=39 · 1.1808^t. R^2=0.974
Austria=61 · 1.3073^t. R^2=0.983
Belgium=88 · 1.2292^t. R^2=0.955
Canada=33 · 1.2468^t. R^2=0.978
France=113 · 1.2538^t. R^2=0.996
Germany=92 · 1.26^t. R^2=0.996
Italy=298 · 1.2154^t. R^2=0.996
×10^4
Japan=77 · 1.0739^t. R^2=0.99
Netherlands=178 · 1.1698^t. R^2=0.998
Norway=70 · 1.2492^t. R^2=0.973
Spain=111 · 1.3468^t. R^2=0.992
Switzerland=110 · 1.2603^t. R^2=0.974
UnitedKingdom=77 · 1.2387^t. R^2=0.988
Growth Rate - Average=23.3225%
Start of Exp. Growth=103 Cases
2020

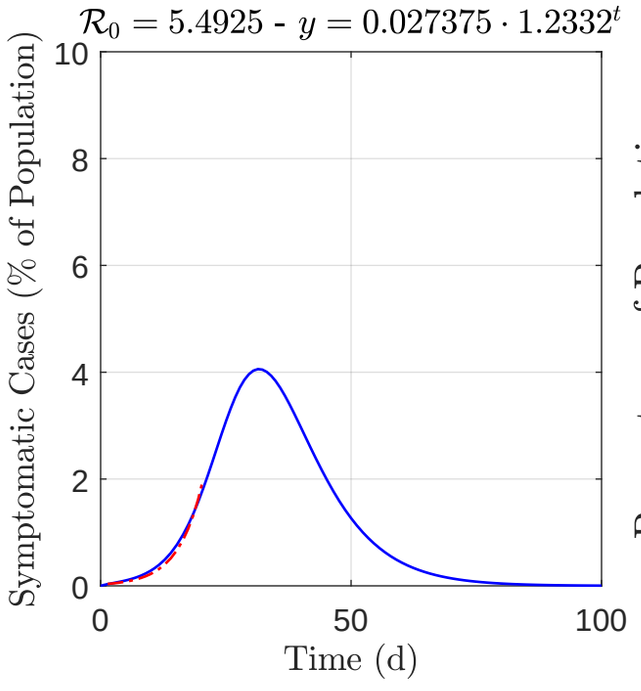

Variation of $\mathcal{R}_0$ vs. λ_{YR}^{-1}, λ_{YR}^{-1}
$\mathcal{R}_0$
24
22
20
18
16
14
12
10
8
6
$\lambda_{YR}^{-1} \in [8,37]$
$\lambda_{AR}^{-1} \in [8,37]$

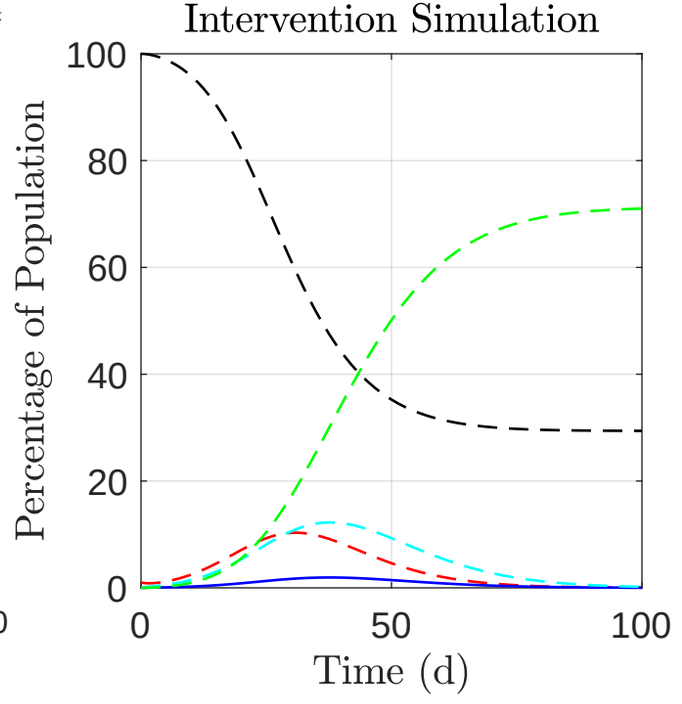

$\lambda_{YR}^{-1} = \lambda_{AR}^{-1} = 37$, $\mathcal{R}_0$ =25.4027
Susceptible
Exposed
Symptomatic
Asymptomatic
Recovered
Percentage of Population
Time (d)


$\mathcal{R}_0 = 25.4027$ - $y = 0.066 \cdot 1.2332^t$
Symptomatic Cases (% of Population)
Time (d)


Intervention Simulation
Percentage of Population
Time (d)


$\lambda_{YR}^{-1} = \lambda_{AR}^{-1} = 8$, $\mathcal{R}_0$ =5.4925
Percentage of Population
Time (d)


$\mathcal{R}_0 = 5.4925$ - $y = 0.027375 \cdot 1.2332^t$
Symptomatic Cases (% of Population)
Time (d)


Intervention Simulation
Percentage of Population
Time (d)