Research article

A closer look at the spreaders of COVID-19 in Wisconsin and the US

Sherry E Scott[1,*]**, Keisha J Cook**[2] **and Kamal Barley**[3]

[1] Department of Mathematics, formerly MSOE, Milwaukee, WI, USA

[2] Department of Mathematics, Tulane University, New Orleans, LA, 70118, USA

[3] Applied Mathematics & Statistics, Stony Brook University, Stony Brook, NY, 11794, USA

* **Correspondence:** Email: sscott@msri.org; Tel: +1-414-445-2487.

Abstract: In this study, we design and use a mathematical model to primarily address the question of who are the main drivers of COVID-19 - the symptomatic infectious or the pre-symptomatic and asymptomatic infectious in the state of Wisconsin and the entire United States. To set the stage, we first briefly simulate and illustrate the benefit of lockdown. With these lockdown scenarios, and in general, the more dominant influence of the the pre-symptomatic and asymptomatic infectious over the symptomatic infectious, is shown in various ways. Numerical simulations for the U.S. show that an increase in testing and isolating for the pre-symptomatic and asymptomatic infectious group has up to 4 times more impact than an increase in testing for the symptomatic infectious in terms of cumulative deaths. An increase in testing for the pre-symptomatic and asymptomatic infectious group also has significantly more impact (on the order of twice as much) on reducing the control reproduction number than testing for symptomatic infectious. Lastly, we use our model to simulate an implementation of a natural herd immunity strategy for the entire U.S. and for the state of Wisconsin (once an epicenter for COVID-19). These simulations provide specific examples confirming that such a strategy requires a significant number of deaths before immunity is achieved, and as such, this strategy is certainly questionable in terms of success.

Keywords: COVID-19, symptomatic infectious, asymptomatic infectious, pre-symptomatic infectious, natural herd immunity

1. Introduction

The coronavirus, known as COVID-19, emerged in Wuhan, China in December 2019, and has since then spread to all countries on earth. By 12/10/2020, the United States reported 15,535,565 cases and 291,403 deaths and although the leading country in both deaths and cases, the U.S. did not implement a national pandemic control initiative. Instead, many states and cities implemented their

own initiatives at various times, and the transmission of COVID-19 has varied from state to state. States also implemented their own testing initiatives. Around 12-10-2020, the reported testing rates for Wisconsin was 562.4 tests per 100,000 people and for the entire US it was 5372 tests per million [1, 2]. With regard to mask usage, the observed mask usage percentage in the US has increased since the beginning of the pandemic. From March 1, 2020 to May 1, 2020, the mask use percentage increased from 5% to 40%. From May 1, 2020 to November 22, 2020, the mask use percentage increased from 71%. Social distancing has also played a role in the transmission of COVID-19. From March 8, 2020 to April 4, 2020, the US experienced a 53% decrease in mobility. From April 4, 2020 to June 24, 2020, the decrease in mobility changed to only -22%. From June 24, 2020 to November 22, 2020, the mobility oscillated between -18% and -25% [3]. The social distancing statistics are due to the fact that different states implemented stay-at-home periods for various intervals of time. All of these factors play a part in the transmission of COVID-19 in the US as a whole, and effectively in each state. Although we do not consider all of these measures directly, our model does include these pertinent features such as mask usage.

Mathematical modelling can help us better understand how COVID-19 cases and mortality is affected based on when and how long lockdown periods are enforced. Modelling can also help us identify the main drivers of the disease and give insight on how best to use the standard control and mitigation strategies. We expand on the standard SIR model, which uses a system of ordinary differential equations to model disease spread through multiple compartments over time. In this study, we do not consider a time-varying system.

Prior studies consider susceptible individuals transitioning to exposed and then to being infectious. We include a pre-symptomatic compartment for individuals who are infectious before the onset of symptoms, if any. We also consider individuals who were tested, resulting in an asymptomatic or symptomatic classification, as well as the population of people who may have self-isolated after being exposed by an infected individual. It is known that over the course of the pandemic, hospitals have been highly stressed and over-crowded due to the influx of COVID-19 patients. In order to study mortality due to the pandemic, we also include compartments that keep track of the number of patients admitted to the hospital and the ICU.

COVID-19 is transmitted by both symptomatic, pre-symptomatic and asymptomatic individuals [4] and most infections show mild or no clinical symptoms [5]. Obtaining real time data separating the two groups is difficult especially since the pre-symptomatic and asymptomatic individuals do not exhibit symptoms and hence have been less likely to be tested and reported; hence, the need for a model. Our study is designed to investigate the role of the symptomatic, pre-symptomatic and asymptomatic more closely using data from the U.S. and the state of Wisconsin which at one time was an epicenter for the COVID-19 pandemic. The research questions are chosen with the aim of analyzing the different aspects of the contribution and effect of asymptomatic and pre-symptomatic transmission versus symptomatic transmission. A better understanding of this transmission is key to developing more informed and effective containment and mitigation strategies for COVID-19 [6] and this information might prove helpful for future coronaviruses and other infectious diseases.

Our model is similar to the model in [7]; however, unlike our model, [7] splits the symptomatic infectious into both mild and severe symptomatic classes and the recovered into tested and not tested. While [7] does include the pre-symptomatic and asymptomatic classes, the role of these infectious classes is not the focus and a comparison of the roles is not done. More specifically, our work builds

on [7] in terms of focusing in more depth (and in various ways) on distinguishing between the contributions of the asymptomatic and pre-symptomatic versus the symptomatic. For example, we include figures that emphasize the contrast in the impact of testing for the two groups and we also align our testing and detection parameters to more so reflect the raw testing rate data.

Asymptomatic and pre-symptomatic transmission has been considered in other studies, e.g., in terms of the data [5, 8] and with an agent-based model [9]. We add to this body of research by considering the role of this silent transmission using a compartmental epidemiological model that is parameterized with mortality data from Wisconsin and the entire U.S. Our model incorporates pertinent aspects of COVID-19 transmission and dynamics and is based on a deterministic system of differential equations. The authors note that the model does not include vaccines since at the time of the development and analysis of the model, vaccines were not yet available, but inclusion of vaccines (and different strains of COVID-19) will be addressed in upcoming work.

We explore scenarios where lockdown is started earlier or extended past the reported lockdown date. We show that both scenarios decrease the number of cases and deaths. We can use these type of simulations to predict future mortality populations and use the information to help with determining when lockdown periods should start and how long the lockdown period should be enforced. As the main goal is to investigate the contribution of the different infectious classes to the spread of the disease, we consider the effects of lockdown in terms of the two infectious groups.

Finally, in response to some public discussions of a natural herd immunity approach, we use our model to run some basic scenarios for implementing such a strategy in the state of Wisconsin and the US as a whole. We simulate the number of deaths that would occur before herd immunity is achieved. The goal is to clearly and concretely illustrate the specific high burden of such an approach in terms of our model and data from Wisconsin and the U.S.

2. Main questions

The overall purpose of this study is to compare and contrast the roles of the asymptomatic and pre-symptomatic infectious individuals in the post-lockdown dynamics of COVID-19. Because both the asymptomatic and pre-symptomatic infectious spread COVID-19 without showing symptoms we group these two compartments together and compare that 'silent transmission" group to the symptomatic infectious. More specifically, we use a mathematical model to address the following questions for Wisconsin and the entire US. The section of the paper in which each question is addressed is given after the question.

- How do changes in lockdown dates (i.e., earlier or later lockdown dates) affect the number of cases - especially in terms of asymptomatic and pre-symptomatic versus symptomatic - as well as cumulative deaths? (Section 4.3.1)
- In general, how does the contribution of the pre-symptomatic and asymptomatic infectious individuals compare to that of the symptomatic infectious in terms of cases and deaths? (Section 4.3.2)
- If we isolate all symptomatic infectious individuals or all asymptomatic and pre-symptomatic infectious individuals, how will this impact the number of deaths and cases? (Section 4.3.3)
- How might a minimum testing rate for pre-symptomatic and asymptomatic compare to that for the symptomatic infectious individuals in order to obtain a reproduction number below 1? (Section

4.3.4)

- How many deaths will occur in a natural herd immunity simulation for Wisconsin and the entire US? (Section 4.4)

2.1. COVID-19

In order to assess and deliver guidelines for individuals to avoid coronavirus infection, the CDC has implemented models equipped with the many stages of the virus [10]. A person who has been infected is considered pre-symptomatic (an individual who is infectious but does not show symptoms at the time of testing) or asymptomatic (a person who is infectious and does not show symptoms throughout the course of the infection). If symptoms set in, the infectious individual is considered symptomatic. The CDC built pandemic planning scenarios varying the infectiousness of asymptomatic individuals compared to symptomatic, varying the percent of asymptomatic individuals, and varying the percentage of transmission of pre-symptomatic individuals. In this paper, we use similar methods to build our model and determine parameter values for the locations and time periods of interest. Figure 1 shows the stages of COVID-19, especially the overlap of the infectious period with the pre-symptomatic period, asymptomatic period and symptomatic period. It should be noted that a pre-symptomatic infectious individual can become either asymptomatic or symptomatic infectious.
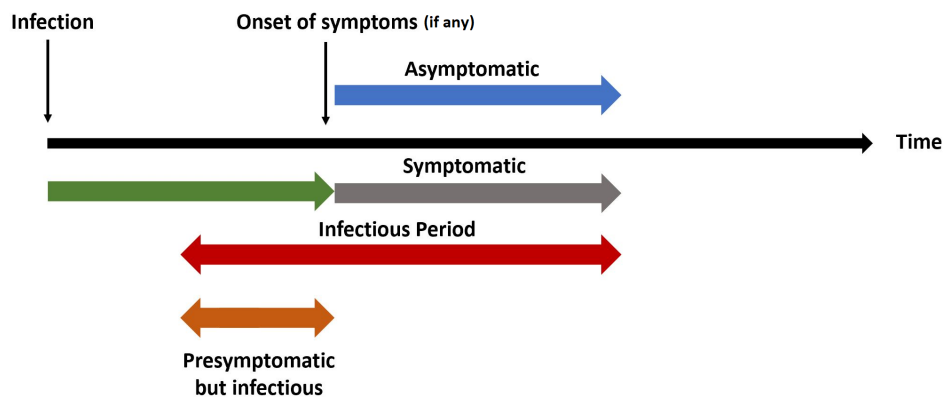

Figure 1. The Stages of COVID-19. Note that the arrow designating "onset of symptoms" is only for individuals who have symptoms, referring to the symptomatic individuals

We are interested in modeling COVID-19 in the United States as a whole, as well as focusing on Wisconsin, since Wisconsin was an epicenter later on in 2020 and is located in the "heartland of America". Table 4 shows some mortality data for Wisconsin and the US.

Table 1. This table gives some mortality data for Wisconsin and the United States during several different time periods.

	Cumulative Deaths			
	Beginning of lockdown	End of lockdown	on August 11, 2020	on November 23, 2020
United States	16767	104803	164519	257779
Wisconsin	10	594	1006	3158

3. Methods

3.1. Model flow diagram

We use a modified Kermack-McKendrik-type epidemic (no human demography) model to better understand the dynamics of COVID-19. Since a main objective of this study is to analyze the role of infectious individuals who show no symptoms, the model Figure 2 includes pre-symptomatic and asymptomatic compartments. We use a deterministic susceptible, exposed, pre-symptomatic, symptomatically-infectious, asymptomatically-infectious, self-isolated, hospitalized, recovered, and ICU patients modeling framework, with the classes denoted as $S(t)$, $E(t)$, $E_P(t)$, $I(t)$, $A(t)$, $J(t)$, $H(t)$, $R(t)$, $C(t)$ respectively; we also include $D(t)$ to track deaths. The model also includes the mitigation/control interventions of face mask usage (with a compliance parameter as well as a face mask efficacy parameter [7]) and testing/detection implementation for each of the infectious classes.
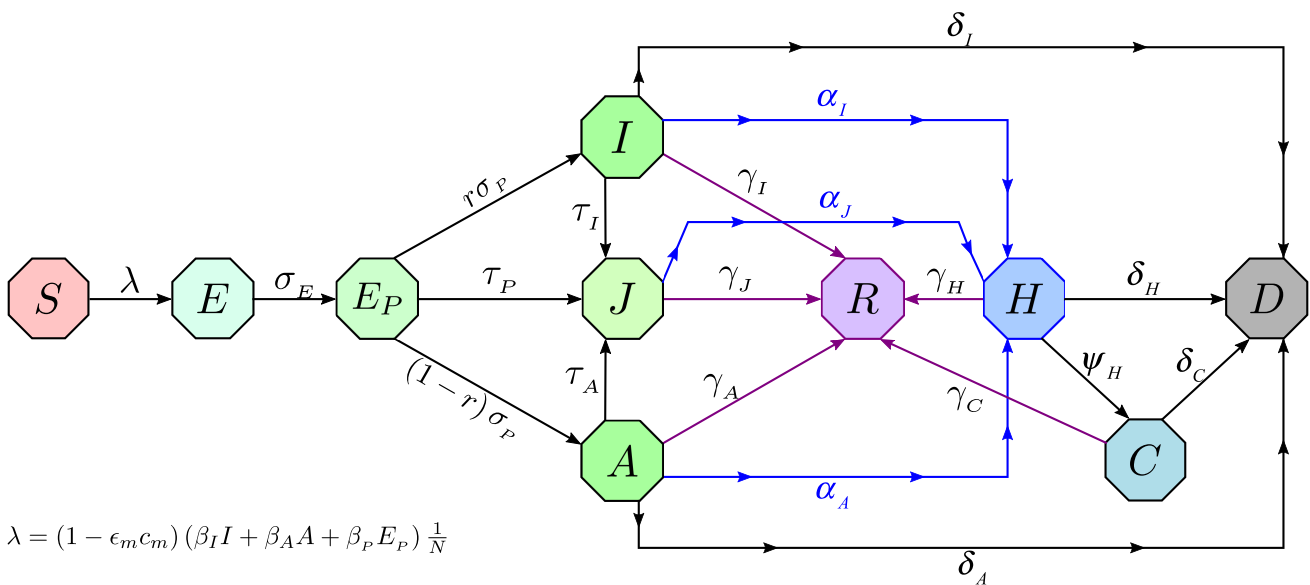

$$\lambda = (1 - \epsilon_m c_m)\left(\beta_I I + \beta_A A + \beta_P E_P\right)\frac{1}{N}$$

Figure 2. 10-compartment flow diagram of the model.

3.2. Data collection

We obtain the observed cumulative deaths and cases data for the state of Wisconsin and the entire US from the COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (2020). The repository contains data beginning at January 22, 2020 (the marked beginning of the pandemic in the US), however, we focus primarily on the data from March 1, 2020 through October 13, 2020. We fit the parameters using mortality data because it has been found to be more reliable than incidence data [7, 11, 12]

The model has 24 parameters and we use values from the literature for 18 of these and estimate the remaining 6 by fitting the model to the observed cumulative mortality data. The parameters that are estimated are as follows: the effective contact rate for the symptomatically-infectious individuals β_I, the effective contact rate for the asymptomatically-infectious individuals β_A, the effective contact rate for pre-symptomatically-infectious individuals β_P, the rate at which presympotomatically-infectious individuals self-isolate τ_P, the rate at which asympotomatically-infectious individuals self-isolate τ_A,

and the hospitalization rate for self-isolating individuals α_J. Parameter fitting was performed using a standard non-linear least-squares approach - i.e. determining the best parameters set that minimizes the sum of the square of the difference between the model outputs for death and the observed values for deaths.

Our model for the transmission dynamics of COVID-19 is given by the following deterministic system of non-linear differential equations.

$$
\begin{aligned}
\frac{dS}{dt} &= -\beta_I(1 - \epsilon_m c_m)\frac{SI}{N} - \beta_A(1 - \epsilon_m c_m)\frac{SA}{N} - \beta_P(1 - \epsilon_m c_m)\frac{SE_P}{N} = -\lambda S \\
\frac{dE}{dt} &= \beta_I(1 - \epsilon_m c_m)\frac{SI}{N} + \beta_A(1 - \epsilon_m c_m)\frac{SA}{N} + \beta_P(1 - \epsilon_m c_m)\frac{SE_P}{N} - \sigma_e E = \lambda S - \sigma_e E \\
\frac{dE_P}{dt} &= \sigma_e E - \sigma_p E_P - \tau_p E_P \\
\frac{dI}{dt} &= r\sigma_p E_P - \alpha_I I - \tau_I I - \gamma_I I - \delta_I I \\
\frac{dA}{dt} &= (1 - r)\sigma_p E_P - \alpha_A A - \tau_A A - \gamma_A A - \delta_A A \\
\frac{dH}{dt} &= \alpha_I I + \alpha_A A + \alpha_J J - \psi_H H - \gamma_H H - \delta_H H \\
\frac{dC}{dt} &= \psi_H H - \gamma_C C - \delta_C C \\
\frac{dJ}{dt} &= \tau_p E_P + \tau_A A + \tau_I I - \alpha_J J - \gamma_J J \\
\frac{dR}{dt} &= \gamma_I I + \gamma_A A + \gamma_C C + \gamma_J J + \gamma_H H \\
\frac{dD}{dt} &= \delta_A A + \delta_I I + \delta_H H + \delta_C C
\end{aligned}
\tag{3.1}
$$

where $N(t) = S + E + EP + I + A + H + J + C + R$ is the total population at time t. The parameter and its description is as follows: ϵ_m - the efficacy of face masks to prevent transmission and acquisition of infection, c_m - the compliance in face mask usage in the community ($0 < c_m \leq 1$), σ - the progression rate from exposed (E) to infectious classes (I or A), r - the fraction of exposed individuals who show clinical symptoms at the end of the incubation period, τ_I- the rate self-quarantined symptomatically-infectious humans self-isolate, $\gamma_I, \gamma_A, \gamma_J, \gamma_H, \gamma_C$- the recovery rates for the subscripted population, ψ_H - the rate of ICU admission for hospitalized individuals, α_I- the hospitalization rate for symptomatically-infectious individuals, α_A- the hospitalization rate for asymptomatically-infectious individuals, δ_I, δ_A, δ_H, δ_C -the disease-induced mortality rates for the subscripted population. We assume that hospitalized individuals do not come in contact with the general population.

The state variables, the parameters for our model and the mask compliance parameters are given in Tables 2, 3, and 4 respectively.

Table 2. State variables description.

State Variable	Description
S	Population of susceptible individuals
E	Population of non-quarantined exposed individuals (infected but not showing symptoms and cannot transmit infection; newly-infected but not infectious)
E_P	Population of pre-symptomatic (infectious) individuals
I	Population of symptomatically-infectious individuals
A	Population of asymptomatically-infectious individuals
J	Population of self-isolated individuals
H	Population of hospitalized individuals
R	Population of recovered individuals
C	Population of individuals in intensive care unit (ICU)
D	Population of COVID-19 deceased individuals

Table 3. Parameter notation, description, values and sources for the model in section 3.1. Some values change depending on the time period and for simulations that are post post-lockdown or involve testing implications, the testing parameter values are chosen to more-so align with reported testing trends.

Parameter	Description	Range	Baseline value	References
β_I	Effective contact rate for symptomatically-infectious individuals	Fitting	Fitting	Fitting
β_A	Effective contact rate for asymptomatically-infectious individuals	Fitting	Fitting	Fitting
β_P	Effective contact rate for pre-symptomatically-infections individuals	Fitting	Fitting	Fitting
ϵ_m	Efficacy of face masks to prevent transmission and acquisition of infection ($0 < \epsilon_m \leq 1$)	0.4 –0.6	0.5	[7]
c_m	Compliance in face mask usage in the community ($0 < c_m \leq 1$)	0.0190-0.1835	per state	[7]
σ_e	Progression rate from exposed (E) to pre-symptomatic infectious class (E_p) infectious class	1/2.5 day^{-1}	1/2.5 day^{-1}	[7]
σ_P	Progression rate from pre-symptomatic infectious class (E_p) to asymptomatically infectious or symptomatically infectious	1/2.5 day^{-1}	1/2.5 day^{-1}	[7]
r	Fraction of exposed individuals who show clinical symptoms at the end of the incubation period	0.4 – 0.6	0.6	[13, 14, 15]
τ_P	Rate at which pre-symptomatic infectious individuals self-isolate	Fitting	Fitting	Fitting
τ_I	Rate self-quarantine symptomatically-infectious humans self-isolate	0.07 - 0.4681 day^{-1}	1/2.5 day^{-1}	[16, 17, 18, 19, 20, 21]
τ_A	Rate at which asymptomatically-infectious humans self-isolate	Fitting	Fitting	Fitting
γ_I	Recovery rate for individuals in the I class	1/30 – 1/3 day^{-1}	1/7 day^{-1}	[13, 14, 15]
γ_A	Recovery rate for individuals in the A class	1/14 – 1/3 day^{-1}	1/7 day^{-1}	[13, 14, 15]
γ_J	Recovery rate for self-isolated individuals	0.0714 – 0.1667 day^{-1}	1/8 day^{-1}	[13]
γ_H	Recovery rate for hospitalized patients	1/30 – 1/3 day^{-1}	1/14 day^{-1}	[13, 14, 15]
γ_c	Recovery rate for ICU patients	0.018 – 0.14 day^{-1}	0.0225 day^{-1}	[13, 22]
ψ_H	Rate of ICU admission for hospitalized individuals	0.02 – 0.1667 day^{-1}	0.083 day^{-1}	[13, 15, 21]
α_I	Hospitalization rate for symptomatically-infectious individuals	0.1111 - 0.3333 day^{-1}	0.2199 day^{-1}	[15, 23, 14]
α_A	Hospitalization rate for asymptomatically-infectious individuals	0.1667 – 0.3333 day^{-1}	1/4 day^{-1}	[24, 13, 25]
α_J	Hospitalization rate for self-isolated individuals	Fitting	Fitting	Fitting
δ_I	Disease-induced mortality rate for symptomatically-infectious individuals	0.001 – 0.1 day^{-1}	0.0225 day^{-1}	[15, 21, 13, 26]
δ_A	Disease-induced mortality rate for asymptomatically-infectious individuals	0.001 – 0.1 day^{-1}	0.0075 day^{-1}	[15, 26]
δ_H	Disease-induced mortality rate for hospitalized individuals	0.001 – 0.1 day^{-1}	0.015 day^{-1}	[14, 15, 21]
δ_c	Disease-induced mortality rate individuals in ICU	0.001 – 0.1 day^{-1}	0.0225 day^{-1}	[14, 15, 26]

Table 4. The mask compliance rates during the respective periods.

	Mask compliance level c_m		
	Pre-lockdown	Lockdown	Post-lockdown
United States	0.0278	0.1835	0.1835
Wisconsin	0.0278	0.1835	0.1392

4. Results

4.1. Computation of reproduction numbers

The basic reproduction number $\mathcal{R}_0$ is the average number of secondary infections produced when one infected individual is introduced into a host population of susceptible individuals. For our model's $\mathcal{R}_0$, we have $\mathcal{R}_0 = \mathcal{R}_{0A} + \mathcal{R}_{0I} + \mathcal{R}_{0P}$ where $\mathcal{R}_{0A}$, $\mathcal{R}_{0I}$, $\mathcal{R}_{0P}$, correspond to the subscripted populations with

$$\mathcal{R}_{0A} = \frac{\beta_A(1-r)}{\alpha_A + \gamma_A + \delta_A}$$

$$\mathcal{R}_{0I} = \frac{\beta_I r}{\alpha_I + \gamma_I + \delta_I}$$

and

$$\mathcal{R}_{0P} = \frac{\beta_P}{\sigma_P}.$$

Note that $\frac{1}{\alpha_A + \gamma_A + \delta_A}$ is the mean infectious duration in the asymptomatic infectious class; so that the $\mathcal{R}_{0A}$ is that number times the product the proportion of exposed individuals that move to the asymptomatic class - $1 - r$ and the asymptomatic effective contact rate - β_A. Similar explanations hold for $\mathcal{R}_{0I}$ and $\mathcal{R}_{0P}$.

The control reproduction number $\mathcal{R}_C$ is the average number of new cases generated by a typical infectious individual introduced into a host population of susceptible individuals with some control measures/interventions in place. Using the next generation operator method and notation found in [27] we have the following computation for the control reproduction number $\mathcal{R}_C$. If we take the column vector $[E \; E_P \; I \; A]$ representing the compartments of infected, then the associated next generation matrices, F and V, for the new infection terms and the transition terms are given respectively as

$$\mathcal{F} = \begin{bmatrix} 0 & (1-\epsilon_m c_m)\beta_P & (1-\epsilon_m c_m)\beta_I & (1-\epsilon_m c_m)\beta_A \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } \mathcal{V} = \begin{bmatrix} \sigma_e & 0 & 0 & 0 \\ -\sigma_e & G_1 & 0 & 0 \\ 0 & -r\sigma_P & G_2 & 0 \\ 0 & -(1-r)\sigma_P & 0 & G_3 \end{bmatrix}$$

where, $G_1 = \sigma_P + \tau_P$, $G_2 = \alpha_I + \gamma_I + \delta_I + \tau_I$, $G_3 = \alpha_A + \gamma_A + \delta_A + \tau_A$. Thus, $\mathcal{R}_C$ is the spectral radius of the next generation matrix given by,

$$\mathcal{R}_C = \rho\left(FV^{-1}\right) = \mathcal{R}_I + \mathcal{R}_A + \mathcal{R}_P$$

where $\mathcal{R}_I = \frac{(1-\epsilon_m c_m)\beta_I r \sigma_P}{G_1 G_2}$, $\mathcal{R}_A = \frac{(1-\epsilon_m c_m)(1-r)\beta_A \sigma_P}{G_1 G_3}$, and $\mathcal{R}_P = \frac{(1-\epsilon_m c_m)\beta_P}{G_1}$. In absence of interventions, i.e. with $\tau_I = \tau_A = \tau_P = \tau_A = \epsilon_m = c_m = 0$, we get the basic reproduction number $\mathcal{R}_0$, as given earlier.

The model has a family of disease-free equilibrium given by: $(S^*, E^*, E_p^*, I^*, A^*, H^*, C^*, J^*, R^*) = (S(0), 0, 0, 0, 0, 0, 0, 0, 0)$, where $S(0)$ is the initial number of susceptible individuals in the population and we obtain the following result.

Theorem 1.
The disease-free equilibrium (DFE) of the model is locally-asymptotically stable if $\mathcal{R}_C < 1$. If $\mathcal{R}_C > 1$, the epidemic rises and then eventually declines to zero.

A basic implication of Theorem 1 is that in order to control the COVID-19 outbreak (to not generate more), it suffices to keep $\mathcal{R}_C < 1$. For epidemic models, (i.e., models without birth and death processes, such as the one that we have), this is a sufficient but not a necessary condition; that is, the disease will eventually die out (in the limit). For models with birth and death processes, this will be a necessary condition.

We can also plot a reproduction number that varies with time; it is called the effective reproduction number [29]. The effective reproduction number is given as $R_c(t) = \frac{(\mathcal{R}_0)(S(t))}{N(t)}$ where $S(t)$ is the susceptible at time t, $N(t) = S + E + E_P + I + A + H + J + C + R$ is the total population at time t and $\mathcal{R}_0$ is the basic reproduction number. Figure 3 illustrates the typical behavior of $R_c(t)$ using our model and data for the U.S.
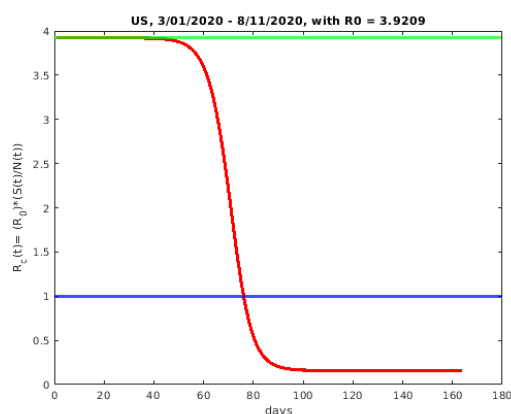
Figure 3. Time varying reproduction number $R_c(t)$ for the U.S.

4.2. Goodness of fit

We measured how well each model fit the data using the T-Test function "t.test" in R. The T-Test measures the difference between two sets of continuous data by comparing two sample means to determine if there is a statistically significant difference between a given model and data. We used a paired T-Test where each vector had the same amount of entries and the values were taken from the same independent variable. We assumed that the data in each vector was normally distributed and that they had approximately equal variances. To test for normality, we used the Shapiro-Wilk Test function "shapiro.test" in R. Let H_0, the null hypothesis, be normally distributed values in the vector. We visualize this using the Q-Q plots in Appendix A, where we assume normality if all the points fall along the reference line. After verifying that the data points are indeed normally distributed, let the null hypothesis for the T-Test be that the difference between the two groups is 0. The T-statistic is defined

as

$$T = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{(s_1^2/n_1) + (s_2^2/n_2)}},$$

where, $\bar{x}_1$ is the mean of the data, $\bar{x}_s$ is the mean of the model values, s_1^2 is the variance of the data, s_2^2 is the variance of the model values, n_1 is the sample size of the data, and n_2 is the sample size of the model values. We conclude that the model for Wisconsin during the lockdown period does indeed fit the data. It is important to note that the results of the T-Test for the United States leads us to believe that the model does not fit the data. However, this is due to the size of the population (the cost of large numbers) in this curve. We say that the model for the US fits the data by eye. Table 5 shows the resulting p-values of the T-Test for Wisconsin and the United States during the lockdown period. The goodness of fit plots can be found in the Appendix (section 6.3) in Figures 16 and 17.

Table 5. The results of the T-Test for the daily and cumulative curves from the beginning of the lockdown period until the end of the lockdown period. The daily column denotes the model vs the data of the daily deaths. The cumulative column denotes the model vs the data of the cumulative deaths.

State	Conclusion	Daily p-value	Cumulative p-value
United States	Different Mean	0.9827	0.1633
Wisconsin	Same Mean	0.0235	2.2e-16

4.3. Sensitivity analysis

We can measure the impact of the sensitivities of the parameters with respect to the control reproduction number. Figure 4 gives sensitivity analysis in terms of the partial rank correlation coefficients (PRCCs) for the parameters [28, 29]. This figure indicates that the parameters that have the most effect on the control reproduction number are the effective contact rate for the pre-symptomatic infectious individuals (β_P), the effective contact rate for the asymptomatic infectious individuals (β_A) and the testing/isolation rate for the pre-symptomatic infectious individuals (τ_P). This result is consistent with our findings that the pre-symptomatic and asymptomatic individuals are the main drivers of the COVID-19 pandemic.
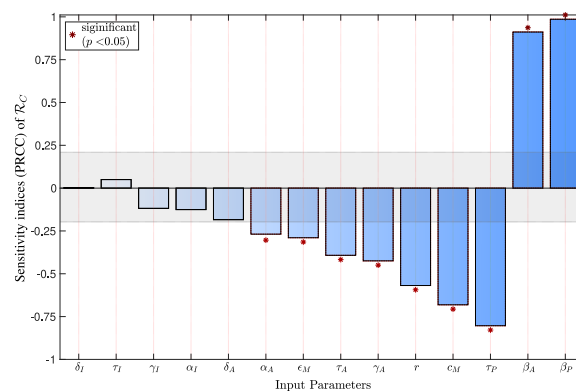

Figure 4. Partial rank correlation coefficients indicating the impact of parameter values on the control reproduction number.

4.4. Simulated epidemics

We fit our model using the cumulative and daily mortality data for Wisconsin and the entire US over the pre-lockdown, lockdown, and post-lockdown periods based on the respective decisions. Wisconsin and the US pre-lockdown period is defined from March 19, 2020 to March 25, 2020 and March 1, 2020 to April 7, 2020 respectively. The lockdown period is defined from March 25, 2020 to May 26, 2020 and April 7, 2020 to May 28, 2020 respectively. For the post-lockdown, we primarily use the dates May 26 to August 11, 2020 and May 28, 2020 to August 11, 2020, respectively [7]. Figure 5 shows the fit for our model with the cumulative deaths data during the pre-lockdown, lockdown and post-lockdown periods.
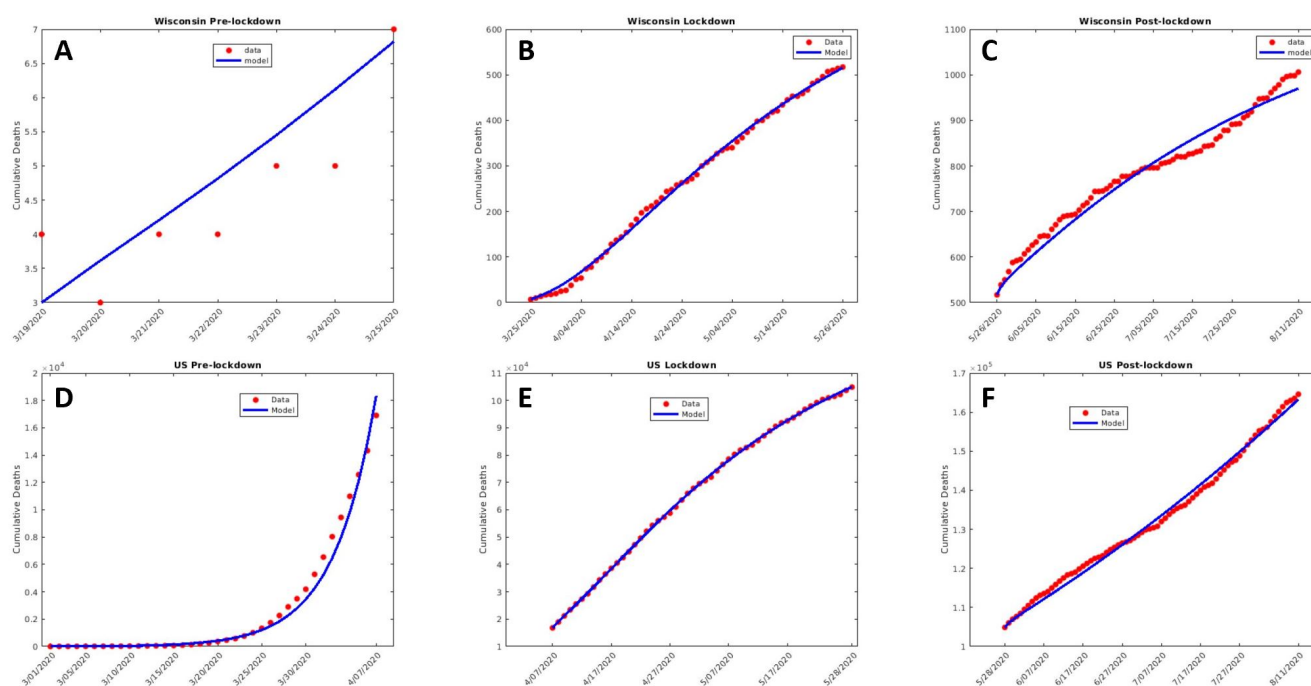
Figure 5. Data fitting of the model for pre-lockdown (A,D), lockdown (B,E) and post-lockdown (C,F), using COVID-19 mortality data for Wisconsin and the US. The plots show the model (blue line) against the data (red dotted line).

4.4.1. Lockdown starting earlier or later

In this section about lockdown we do the following: (1) illustrate the effects of earlier lockdowns on the number of cases of asymptomatic and pre-symptomatic versus symptomatic, (2) show the benefit of lockdown extensions for the state of Wisconsin, and (3) give simulations of how lockdown might have helped during/before the winter holiday period.

The first lockdown was implemented in the month of March. Since an earlier lockdown would have most likely prevented the disease from gaining a foothold in these areas, an earlier lockdown reduces the number of cumulative deaths. In general, an earlier lockdown also results in a decrease in daily cases for both the symptomatic and asymptomatic/pre-symptomatic compartments, and moreover, during these earlier lockdowns, the asymptomatic and pre-symptomatic dominate by a factor of at least 2,

as shown in the daily cases plots for the entire US in Figure 6. A decrease occurs each time we add one week to the beginning of the lockdown period, e.g., the maximum daily cases for the pre-symptomatic and asymptomatic goes from about 16,000 (Fig 6A) to 6000 (Fig 6B).
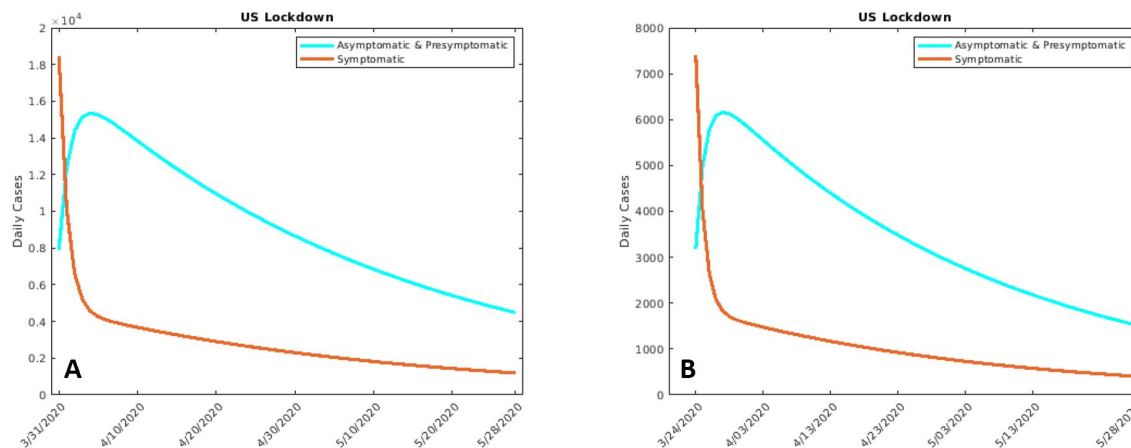

Figure 6. Daily cases for the asymptomatic/pre-symptomatic individuals (cyan curve) versus the symptomatic individuals (red curve) for the entire US, if (A) lockdown began 1 week earlier on March 31, 2020 and if (B) lockdown began 2 weeks earlier on March 24, 2020.

A different scenario involves extending lockdown. If for Wisconsin, we use the original lockdown start dates but extend the lockdown periods by 2 and 4 weeks - Figure(7), we see a decrease in cumulative deaths of about a 14% at the end under a 4 week extension, as shown in Figure 7B.
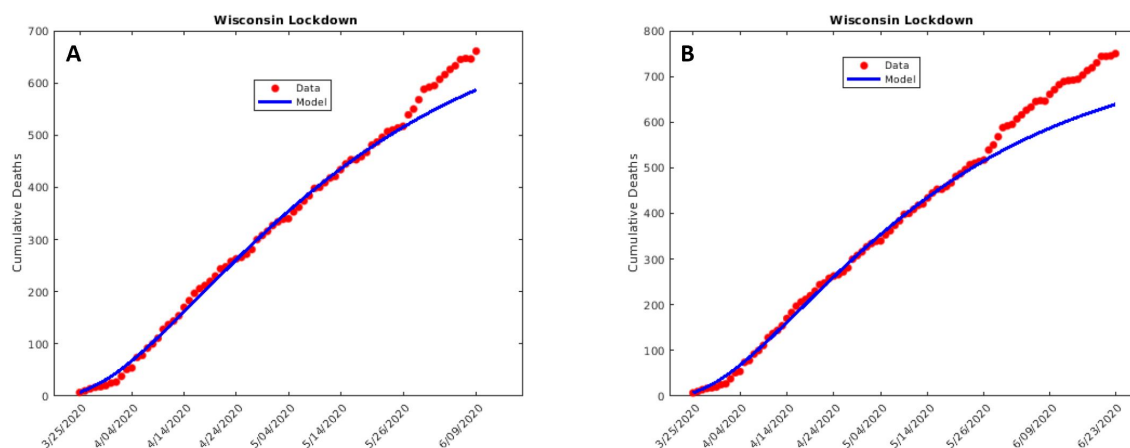
Figure 7. The cumulative deaths for Wisconsin if (A) lockdown was lifted 2 weeks later ending on June 9, 2020 and if (B) lockdown was lifted 4 weeks later ending on June 23, 2020.

These varying lockdown dates scenarios can be used to assess the efficacy of lockdown periods and to guide the manner in which the lockdown periods are implemented. For example, one other interesting lockdown scenario is to consider what would happen if the entire U.S. and Wisconsin implemented a lockdown during the month of December 2020 for two weeks. In order to simulate such a strategy

we reduce the contact rate parameters - i.e., the betas - by 40% [21] and run the simulations forward. As depicted in Figure 8 below, the simulations give that such a lockdown would result in a marked decrease in the number of deaths - around 10% decrease for the U.S. (Fig 8A).
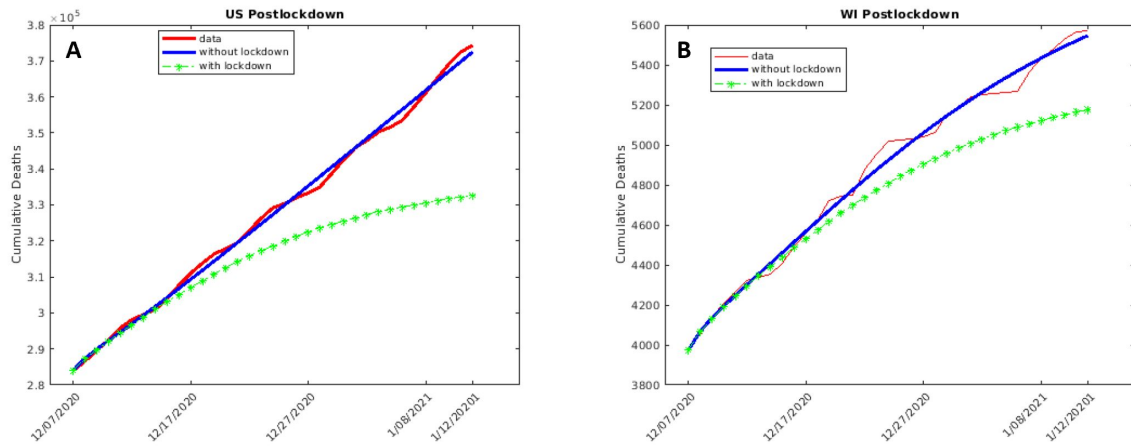

Figure 8. Cumulative deaths in (A) the U.S. and (B) Wisconsin. The green line represents lockdown implementation in December, the blue line represents no lockdown implementation, and the red line represents the data from John Hopkins.

4.4.2. Simulations for Asymptomatic, pre-symptomatic, and Symptomatic

We now focus more on comparing the contributions of the symptomatic infectious group with the asymptomatic and pre-symptomatic infectious group. First, we consider plotting the daily cases for the two groups against each other during the postlockdown period ending on 8/11/2020. In both the state of Wisconsin and the entire U.S., we find that the number of daily cases for the asymptomatic and pre-symptomatic is at least twice more than the daily cases for the symptomatic - Figure 9. We can
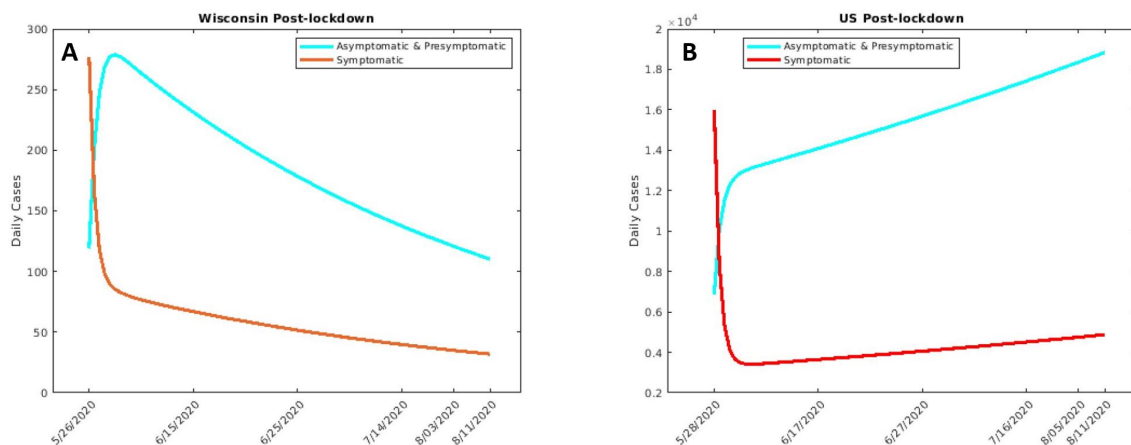
Figure 9. The plots show the daily cases for pre-symptomatic and asymptomatic versus symptomatic for (A) the Wisconsin post-lockdown period from May 26, 2020 to August 11, 2020 and (B) the U.S. post-lockdown period from May 28, 2020 to August 11, 2020.

also focus on each group separately. The parameter β_I represents the contacts by the symptomatically

infectious individuals that could transmit the disease. By setting $\beta_I = 0$, we eliminate that transmission and focus on the effect of the pre-symptomatic and asymptomatic individuals. Similarly, fixing $\beta_A = \beta_P = 0$, targets the effect of the symptomatically infectious. In order to further discern the role of the asymptomatic and pre-symptomatic infectious versus that of the symptomatic infectious, we consider these zero contact rate scenarios, in terms of the number of daily cases Wisconsin and the entire US. In Figure 10, we show these results for post-lockdown periods and we again get that the larger contribution comes from the asymptomatic and pre-symptomatic individuals.
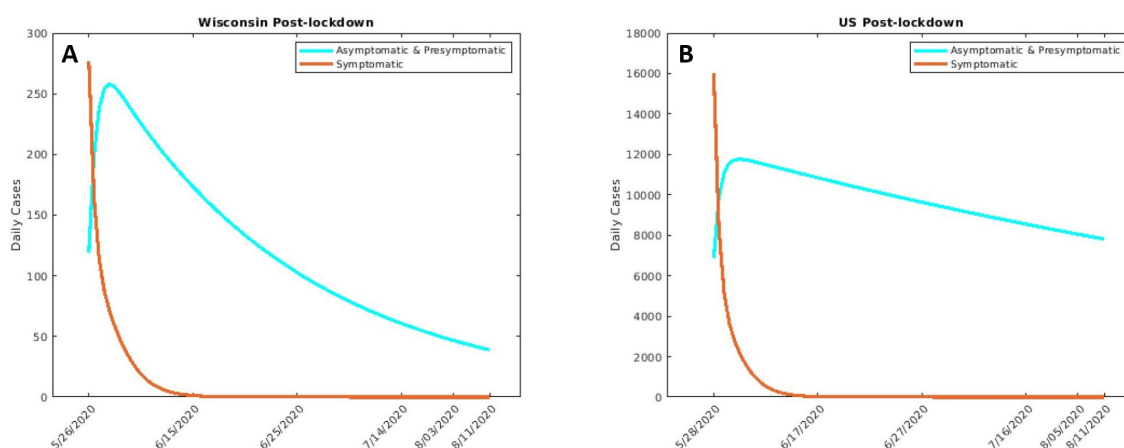
Figure 10. The plots show the daily cases with $\beta_I = 0$ (The cyan curve focuses on the impact of asymptomatic/pre-symptomatic individuals) and with $\beta_A = \beta_P = 0$ (The red curve focuses on the impact of symptomatic individuals) during (A) the Wisconsin post-lockdown period from May 26, 2020 to August 11, 2020, and (B) the U.S. post-lockdown period from May 28, 2020 to August 11, 2020.

In Figure 11, we again plot the $\beta_I = 0$ (green curve) versus the $\beta_A = \beta_P = 0$ (magenta curve) scenario, but here we consider the effect in terms of cumulative cases and cumulative deaths. The solid blue curve corresponds to the model using baseline β values. Both types of plots depict that the main drivers are the asymptomatic and pre-symptomatic individuals and for cumulative deaths the difference is around a factor of 1.3.

4.4.3. Isolating Symptomatic and Asymptomatic/Pre-symptomatic

Recall that the τ values give the rate for testing/detection of infected individuals and subsequent self-isolation. In this section we compare and contrast the effect of isolating the symptomatic against the effect of isolating the asymptomatic and pre-symptomatic by changing the appropriate τ values. For example, to address the effect of isolating symptomatic infectious individuals, we can increase the value of τ_I (the rate that symptomatically infectious individuals self-isolate) while keeping all other rates the same. We consider this scenario for the entire US and find that when τ_I increases (so that the number of symptomatically infectious individuals that are self-isolated and taken out of the general population increases), the number of cumulative deaths decreases by as much as 12.5% (Fig 12A). Alternatively, if the rate of self-isolation for the asymptomatic and pre-symptomatic infectious individuals - τ_A and τ_P respectively - are increased and the τ_I is kept at the baseline value for the US, the number of cumulative deaths decreases by about 30% at the end of the period - see Figure
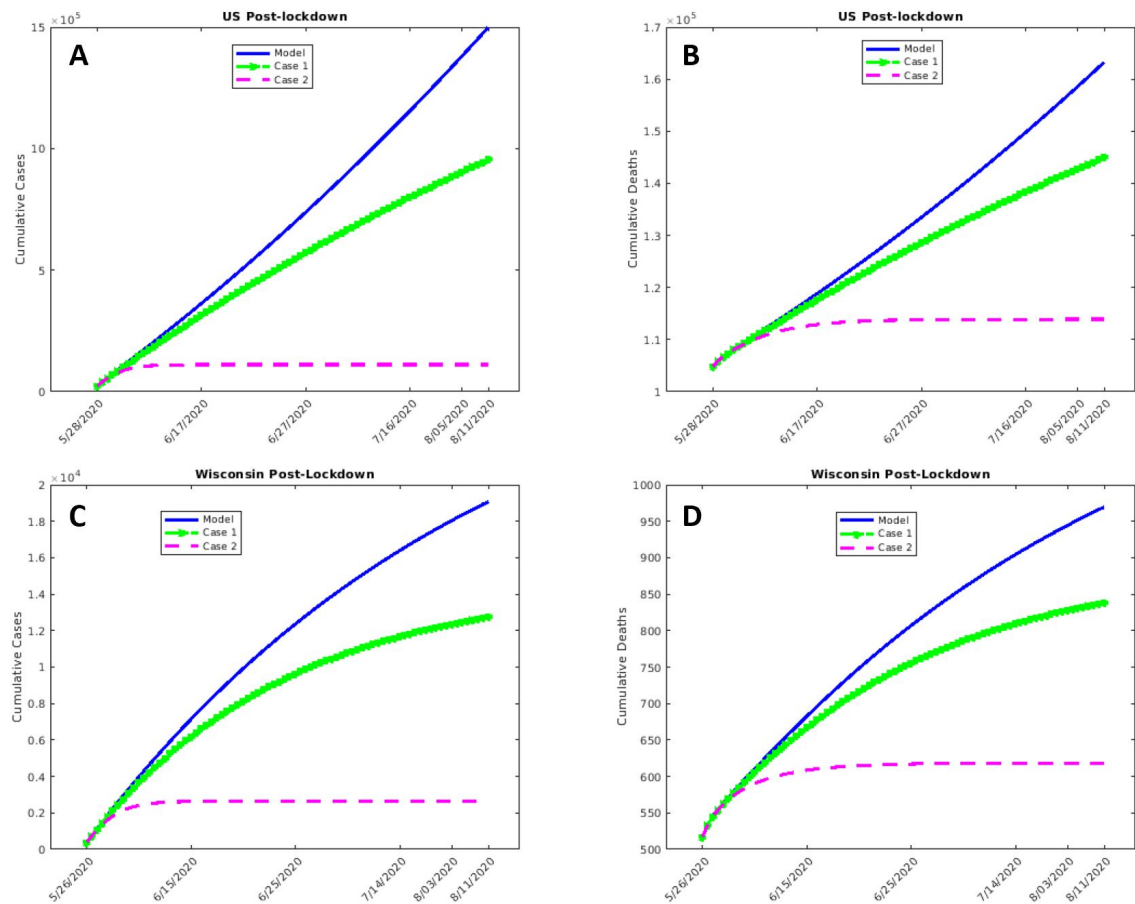
Figure 11. The cumulative cases (A,C) and cumulative deaths (B,D) over the respective post-lockdown periods where **Case 1** corresponds to $\beta_I = 0$ (The green curve focuses on the impact of asymptomatic/pre-symptomatic individuals). **Case 2** corresponds to $\beta_A = \beta_P = 0$ (The magenta curve focuses on the impact of symptomatic individuals).

12B. Thus, the decrease is more pronounced (around twice as much) when the τ_A and τ_P are increased than when the τ_I is increased which indicates that the asymptomatic and pre-symptomatic have more impact. Figure 13A focuses on the difference in impact of a 50% increase in testing and shows that the asymptomatic and pre-symptomatic testing gives up to about four times as much percent decrease in deaths as that of the symptomatic. We find that a 50% increase in τ_A and τ_P, results in roughly a 20% decrease in cumulative deaths for Wisconsin (Fig 13B). Note that all of the results confirm the greater influence of the asymptomatic and pre-symptomatic infectious.

4.4.4. Minimum testing needed

A key part of controlling COVID-19 involves controlling the reproduction number. With this in mind, we consider the question concerning the minimum rate of testing/detection required to lower the control reproduction number, $\mathcal{R}_C$ below one. This question is more difficult to answer directly, but we can begin to get at an answer by again considering the effect of increasing testing. For example, for the entire US from May 28 to Aug 11 (after lifting of lockdown) if the testing rate for
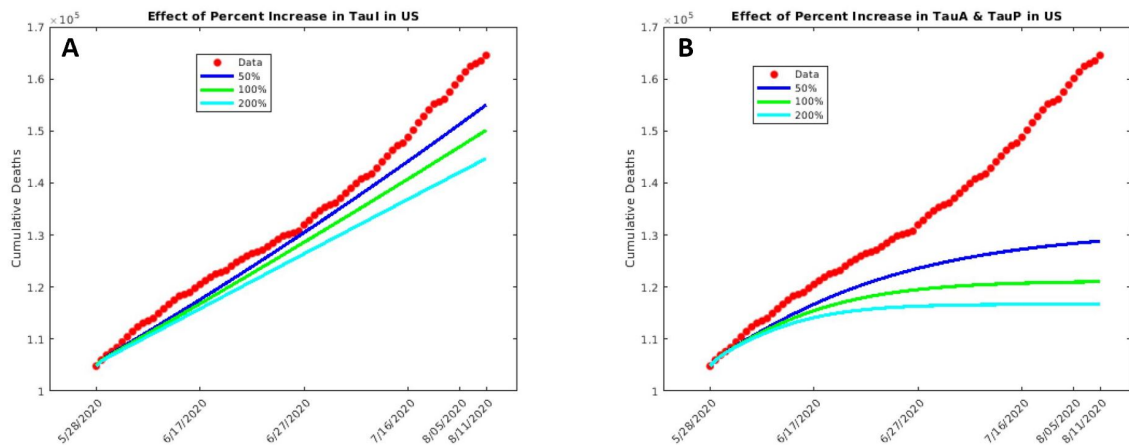
Figure 12. The effect of an increase in testing of symptomatic individuals versus asymptomatic and pre-symptomatic individuals on the cumulative deaths in the US during the post-lockdown period from May 28, 2020 to August 11, 2020. In both plots, the red dotted line represents the data, the blue line represents a 50% increase, the green line represents a 100% increase, and the cyan line represents a 200% increase. (A) The plot describes the increase in τ_I, symptomatic individuals. (B) The plot describes the increase in τ_A and τ_P, asymptomatic and pre-symptomatic individuals.
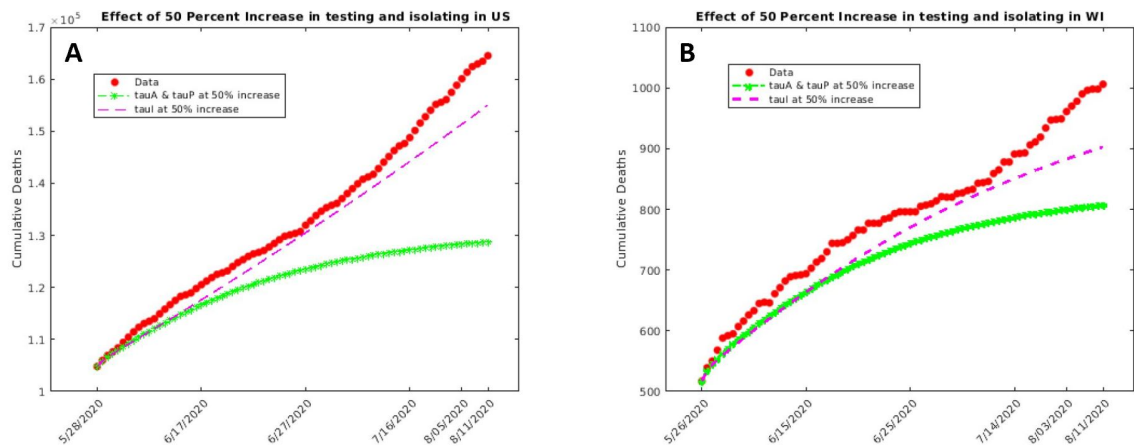
Figure 13. The effect of a 50% increase in testing of asymptomatic and pre-symptomatic versus symptomatic on the cumulative deaths. In both scenarios, the red dotted line represents the data, the green dot-dashed line represents the model of the cumulative deaths by increasing τ_A and τ_p (asymptomatic and pre-symptomatic individuals) by 50%, and the purple dashed line represents the model of the cumulative deaths by increasing τ_I (symptomatic individuals) by 50%. (A) The US post-lockdown period from May 28, 2020 to August 11, 2020. (B) The Wisconsin post-lockdown period from May 26, 2020 to August 11, 2020.

asymptomatic and pre-symptomatic is 0.2 - that is, if τ_A and τ_P are both set to 0.2, while $\tau_I = 0.4$, then the reproduction control number $\mathcal{R}_C$ is 1.0346 (so just above 1). If τ_A and τ_P are bumped above 0.2, for example to 0.25, while τ_I remains at 0.4, then $\mathcal{R}_C$ falls below 1 - to 0.95178 - and the resulting

decrease in cumulative death (Fig 14A) and daily cases (Fig 14B) is given in Figure 14. To get a better appreciation of the impact of testing for asymptomatic and pre-symptomatic cases, note that if τ_A and τ_P are kept at their baseline fitted values, then τ_I must be increased to 1.7 (from a baseline value of 0.4) to get $\mathcal{R}_C = 0.9945$. Of course, the challenge with all of these testing scenarios is that the testing of asymptomatic and pre-symptomatic individuals is more difficult since those individuals do not have symptoms, so in order to achieve these testing rates, widespread and frequent testing should be implemented.
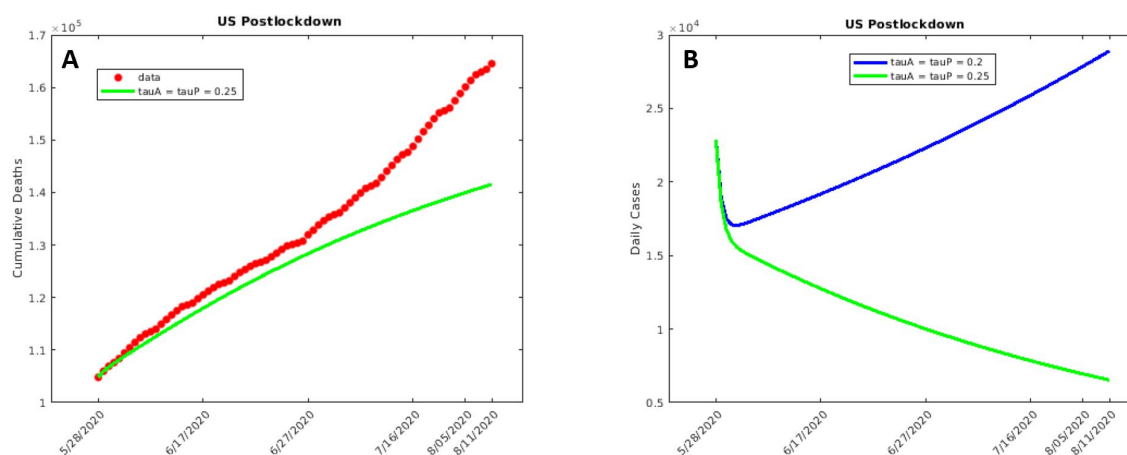
Figure 14. The effect of an increase in τ_A and τ_P (asymptomatic and pre-symptomatic individuals) during the post-lockdown period (May 28, 2020 to August 11, 2020) in the US. (A) The red dotted line represents the data from John Hopkins. The green line represents the model of the cumulative deaths given $\tau_A = \tau_P = 0.25$. (B) The blue line represents the model of the daily cases given $\tau_A = \tau_P = 0.2$. The green line represents the model of the daily cases given $\tau_A = \tau_P = 0.25$.

4.5. Forward simulation of the pandemic

At one point in 2020, as the U.S. and world leaders grappled with the pandemic, the idea of a natural herd immunity approach was widely discusses and even endorsed. Using out model and the data for Wisconsin and the U.S., we provide specific simulations that show that a natural herd immunity approach would cost many lives and as such is, at the very least, a rather dubious approach. Natural herd immunity refers to the process by which enough of the population achieves immunity via natural recovery from the disease, so that the remaining population that is not immune also receives protection against the acquisition of the disease. The process involves a natural herd immunity threshold, that is, the minimum number of people required to achieve disease-acquired immunity. For a disease with a basic reproduction number of $\mathcal{R}_0$, the necessary minimum fraction of people who must achieve immunity can be given as $1 - 1/\mathcal{R}_0$. We simulate for both Wisconsin (a state which was considered an epicenter later in 2020) and the entire US, a scenario in which COVID-19 is allowed to run its natural course (with no implementation of any new control or mitigation strategies) until natural herd immunity is achieved. For both Wisconsin and the entire US, the number of deaths that occur before

reaching the herd immunity threshold, is considerably high. For the state of Wisconsin, the number of deaths that occurs is approximately 80,000 (Fig 15A) and for the entire U.S., the number of deaths is more than 5 million (Figure 15B).
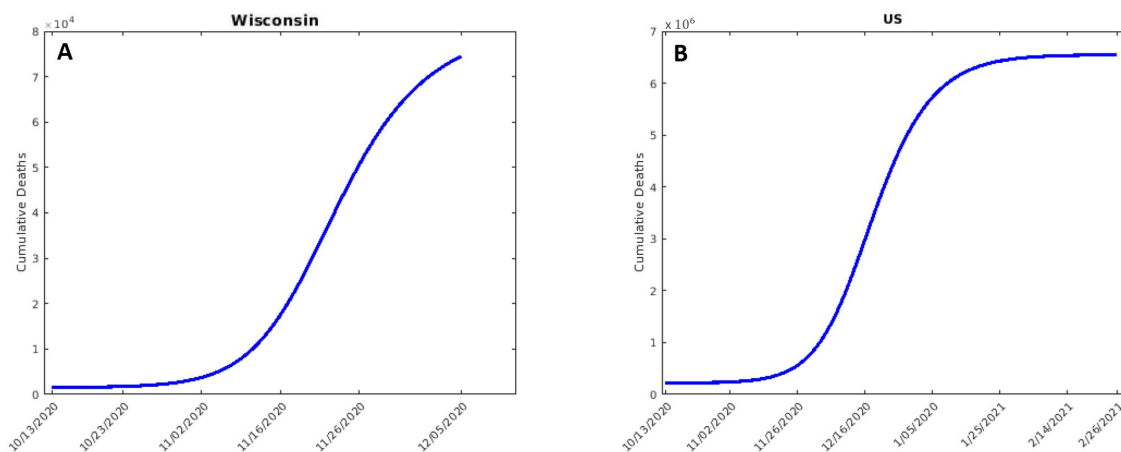

Figure 15. The number of deaths that will occur before the natural herd immunity threshold is achieved (A) Wisconsin simulations from October 13, 2020 to December 5, 2020. (B) The US simulations from October 13, 2020 to February 26, 2021.

5. Discussion

From various perspectives, our analysis indicates that overall, the drivers of COVID-19 are the asymptomatic and pre-symptomatic infectious individuals, i.e., the silent transmission. Even with vaccines, it is critical to control the spread especially as more and more mutations arise, so gaining a better understanding of containment and mitigation strategies remains key for dealing with COVID-19. Our work emphasizes that these strategies must consider the silent transmission. In particular, the respective answers to the research questions posed in Section 2 are as follows: (1) Our simulations show that even under lockdown, the asymptomatic and pre-symptomatic infectious drive the spread. And although lockdown might be difficult in terms of economic impact, the benefit of even a short lockdown (two weeks) in reducing deaths is considerable - around a 10% decrease for the U.S. in the December 2020 simulations (2) More cases are due to asymptomatic and pre-symptomatic infectious than the symptomatic infectious and the daily cases due to the silent transmission is at least twice the number of symptomatic infectious daily cases. If the effective contact rate for the symptomatic infectious is set to zero (so that the focus is on the silent spread), then the asymptomatic and pre-symptomatic cumulative cases are more than five times that of the symptomatic infectious and asymptomatic and pre-symptomatic individuals account for more than 1.3 times as many of the cumulative deaths. (3) We see that an increase in testing for the asymptomatic and pre-symptomatic infectious has a greater impact than the symptomatic infectious and that (4) to lower the control reproduction number below 1, requires a significantly greater increase in testing/isolating of symptomatic infectious than for asymptomatic and pre-symptomatic. Lastly, (5) we provide two simulations to illustrate and emphasize that a notion of natural herd immunity approach with COVID-19 extracts a heavy burden due to the high

number of deaths required to achieve the herd immunity threshold.

For policy makers, our results show that the dominant role of the asymptomatic and pre-symptomatic in terms of the number of cases and deaths and the testing impact, implies that consideration of this silent transmission is key to controlling the spread of the disease. This is especially true after holiday periods and large gathering events which are often notorious for reduced and lax implementation of control and mitigation strategies such as social distancing and mask usage. Thus, in the schools and workplaces, testing should be intentionally and systematically increased after such occurrences in order to control the silent transmission by the newly infectious individuals who do not (yet) have symptoms. Perhaps this work provides some support for the development of home administered testing. Our simulations also show that although lockdown might be difficult in terms of economic impact, the benefit of even a short (perhaps limited or partial) lockdown in reducing deaths and cases is significant enough to justify the option at least during high risk situations. In particular, the asymptomatic and pre-symptomatic infectious must be definitively accounted for as policy makers move forward.

Acknowledgements

Conflict of interest

The authors declare no conflicts of interest.

References

1. *COVID-19 Dashboard by the Center for Systems Science and Engineering*, 2020. Available from: `https://coronavirus.jhu.edu/map.html`.

2. H. Ritchie, E. Ortiz-Ospina, D. Beltekian, E. Mathieu, J. Hasell, B.Macdonald, et al., Coronavirus Pandemic (COVID-19), *Our World in Data*, (2020), Available from: `https://ourworldindata.org/coronavirus`.

3. *COVID-19 Statistics*, 2020. Available from: `https://covid19.healthdata.org/united-states-of-america?view=mask-use&tab=trendl`.

4. Y. Bai, L. S. Yao, T. Wei, F. Tian, D. Y. Jin, L. J. Chen, et al., Presumed asymptomatic carrier transmission of COVID-19, *JAMA*, 2020.

5. World Health Organization, Coronavirus disease 2019 (COVID-19): Situation report, 46, *WHO*, 2020.

6. C. Savvides, R. Siegel, Asymptomatic and presymptomatic transmission of SARS-CoV-2: A systematic review, preprint,`arXiv:PMID32587980`.

7. C. N. Ngonghala, E. A. Iboi, A. B. Gumel, Could masks curtail the post-lockdown resurgence of COVID-19 in the US?, *Math. Biosci.*, **329** (2020), 1–18.

8. L. C. Tindale, J. E. Stockdale, M. Coombe, E. S. Garlock, W. Y. Venus Lau, M. Saraswat, et al., Evidence for transmission of COVID-19 prior to symptom onset, *eLife*, **09** (2020).

9. S. M. Moghadas, M.C. Fitzpatrick, P. Sah, A. Pandey, A. Shoukat, B.H. Singer, et al., The implications of silent transmission for the control of COVID-19 outbreaks *PNAS*, 2020.

10. *Coronavirus*, 2020. Available from: `https://www.cdc.gov/coronavirus/2019-84ncov/hcp/planning-scenarios.html`.

11. H. Lau, T. Khosrawipour, P. Kocbach, H. Ichii, J. Bania, V. Khosrawipour, Evaluating the massive underreporting and undertesting of COVID-19 cases in multiple global epicenters, *Pulmonology*,(2020).

12. L. H. Sun, J. Achenbach , CDC chief says coronavirus cases may be 10 times higher than reported, *The Washington Post (Health)*, (July 3, 2020).

13. C. N. Ngonghala, E. Iboi, S. Eikenberry, M. Scotch, C. R. MacIntyre, M. H. Bonds, et al., Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus, *Math. Biosci.*, **325** (2020).

14. A. Gumel, E. Iboi, C. Ngonghala, Using Mathematics to Understand COVID-19 Dynamics, preprint *submitted to Notic. of the Am. Math. Soc.*, (2020).

15. S. Eikenberry, M. Mancuso, E. Iboi, T. Phan, K. Eikenberry, Y. Kuang, et al., To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic, *Infect. Dis. Model.*, **5** (2020), 293–308.

16. L. Worden, R. Wannier, S. Blumberg, A. Ge, G. Rutherford, T. Porco, Estimation of effects of contact tracing and mask adoption on COVID-19 transmission in San Francisco: a modeling study, preprint,`arXiv:PMID32577672`.

17. G. Chowell, D. Chowell, K. Roosa, R. Dhillon, D. Srikrishna, Sustainable social distancing through facemask use and testing during the Covid-19 pandemic, preprint, `arXiv:medRxiv 2020.04.01.20049981`.

18. T. P. Sheahan, A. C. Sims, S. Zhou, R. L. Graham, A. J. Pruijissers, M. L. Agostini, et al., An orally bioavailable broad-spectrum antiviral inhibits SARS-CoV-2 in human airway epithelial cell cultures and multiple coronaviruses in mice, *Sci. Transl. Med.*, **12** (2020).

19. A. J. Kucharski, P. Klepac, A. J. K. Conlan, S. M. Kissler, M. L. Tang, H. Fry, et al., Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study, *Lancet Infect. Dis.*, **20** (2020), 1151–1160.

20. S. Tian, Z. Chang, Y. Wang, M. Wu, W. Zhang, G. Zhou, et al., Clinical Characteristics and Reasons for Differences in Duration From Symptom Onset to Release From Quarantine Among Patients With COVID-19 in Liaocheng, China, *Front. in Med.*, **7** (2020).

21. E. A. Iboi, C. N. Ngonghala, A. B. Gumel, Will an imperfect vaccine curtail the COVID-19 pandemic in the U.S.? *Infect. Dis. Model.*, **5** (2020), 510–524.

22. J. Gold, K. K. Wong, C. M. Szablewski, P. R. Patel, J. Rossow, J. da Silva, et al., Characteristics and Clinical Outcomes of Adult Patients Hospitalized with COVID-19 -Georgia, March 2020, *Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report*, **69** (2020), 545–550.

23. S. Garg, L. Kim, M. Whitaker, A. O'Halloran, C. Cummings, R. Holstein, et al., Hospitalization Rates and Characteristics of Patients Hospitalized with Laboratory-Confirmed Coronavirus Disease 2019 - COVID-NET, 14 States, March 1 - 30, 2020, *Centers for Disease Control and Prevention Morbidity and Mortality Weekly Report*, **69** (2020), 458–464.

24. *Management of Patients with Confirmed 2019-nCoV*, 2020. Available from: `https://www.cdc.gov/coronavirus/2019-ncov/hcp/clinical-guidance-management-patients.html`

25. Y. Wang, D. Zhang, G. Du, R. Du, J. Zhao, Y. Jin, et al., Remdesivir in adult with severe COVID-19: a randomised, double-blind, placebo-controlled, multicentre trial, *The Lancet*, **395** (2020), 1569–1578.

26. N. M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, et al., Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand, *Imp. Coll. Lon.*, 2020.

27. P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48..

28. S. Marino, I. Hogue, C. Ray, D. Kirschner, A Methodology For Performing Global Uncertainty and Sensitivity Analysis in Systems Biology, *J. Theor. Biol.*, **254** (2008), 178–196.

29. A. Gumel, E. Iboi, C. Ngonghala, E. Elbasha, A primer on using mathematics to understand COVID-19 dynamics: Modeling, analysis and simulations, *Infect. Dis. Model.*, **6** (2021), 148–168.

Appendix A

Goodness of Fit

In this section, we show the plots for the goodness of fit of the lockdown period for the US and Wisconsin. We also display the boxplot of the values for each set in order to show the mean of the different sets of values.
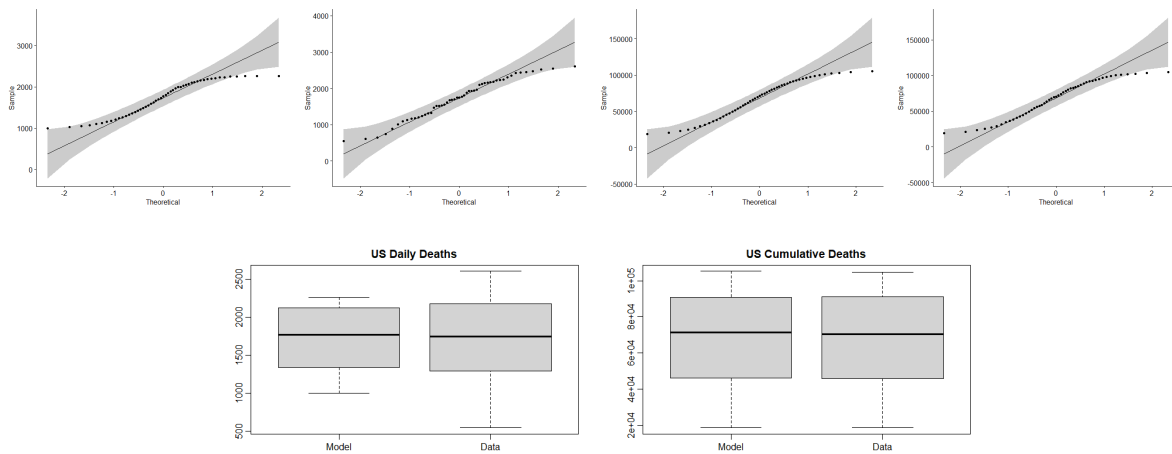
Figure 16. US **Row 1**: Q-Q plots for the (1) Daily death values from the model (2) Daily death values from the data (3) Cumulative death values from the model (4) Cumulative death values from the data. **Row 2**: Box plot of the (1) Daily death values from the model and data (2) Cumulative death values from the model and data.
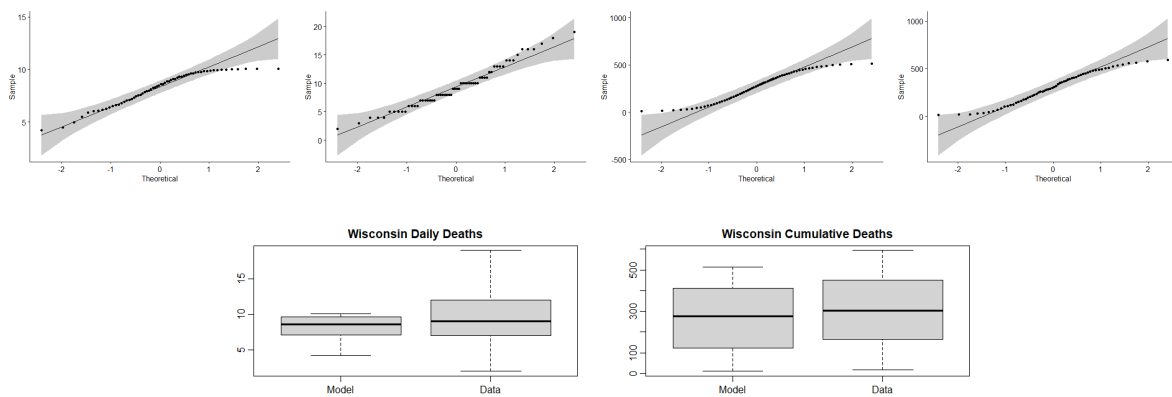

Figure 17. Wisconsin **Row 1**: Q-Q plots for the (1) Daily death values from the model (2) Daily death values from the data (3) Cumulative death values from the model (4) Cumulative death values from the data. **Row 2**: Box plot of the (1) Daily death values from the model and data (2) Cumulative death values from the model and data.