

A NEW APPROACH TO THE DYNAMIC MODELING OF AN INFECTIOUS DISEASE

B. SHAYAK^{1,*} AND MOHIT M. SHARMA²

Abstract. In this work we propose a delay differential equation as a lumped parameter or compartmental infectious disease model featuring high descriptive and predictive capability, extremely high adaptability and low computational requirement. Whereas the model has been developed in the context of COVID-19, it is general enough to be applicable with such changes as necessary to other diseases as well. Our fundamental modeling philosophy consists of a decoupling of public health intervention effects, immune response effects and intrinsic infection properties into separate terms. All parameters in the model are directly related to the disease and its management; we can measure or calculate their values a priori basis our knowledge of the phenomena involved, instead of having to extrapolate them from solution curves. Our model can accurately predict the effects of applying or withdrawing interventions, individually or in combination, and can quickly accommodate any newly released information regarding, for example, the infection properties and the immune response to an emerging infectious disease. After demonstrating that the baseline model can successfully explain the COVID-19 case trajectories observed all over the world, we systematically show how the model can be expanded to account for heterogeneous transmissibility, detailed contact tracing drives, mass testing endeavours and immune responses featuring different combinations of temporary sterilizing immunity, severity-reducing immunity and antibody dependent enhancement.

Mathematics Subject Classification. 34K05, 92B05.

Received November 17, 2020. Accepted May 2, 2021.

1. INTRODUCTION

We split this discussion into two Sections, one general and the second more specific.

1.1. Overview of disease modeling approaches

With the spread of COVID-19 like wildfire all over the world, infectious disease dynamics has suddenly been promoted from a niche area of dynamical systems theory to the foremost topic in applied mathematics and sciences research. Mathematical modeling is one of the widely used scientific tools which allows us to predict

Keywords and phrases: COVID-19, delay differential equation, phenomena-driven parameters, low computational cost, public health interventions, immune response.

¹ Theoretical and Applied Mechanics, Sibley School of Mechanical and Aerospace Engineering, Cornell University, Ithaca 14853, NY, USA.

² Population Health Sciences, Weill Cornell Medicine, 1300 York Avenue, New York City 10065, NY, USA.

* Corresponding author: sb2344@cornell.edu

the trajectories of the disease in advance and take intervention measures accordingly. There are four approaches to modeling of an infectious disease, which we describe below.

- Lumped parameter or compartmental model: These are deterministic differential equation models, like S-I-R and S-E-I-R. In most cases these models use ordinary differential equations (ODE), some though not many use delay differential equations (DDE) and a few use partial differential equations (PDE). The advantage of these models is that they are physically insightful and computationally tractable (especially ODE and DDE models). The drawback is that the division of population into compartments automatically requires averagings and assumptions; the extent to which these limitations fetter the performance of the model depends to a large degree on the model itself. We shall elaborate on this type of model in the next Section.
- Agent-based model: This model considers people as lattice sites on a network. Each site can be in one of several states – typical states are healthy and susceptible, exposed and non-infectious, asymptomatic infectious and at large, symptomatic and quarantined etc. A lattice site contracts the infection with a certain user-defined probability if its one or more neighbours are infectious, and then progresses through the successive states with user-defined probabilities and durations. The advantage of this model is that it is the closest representation of reality and hence is capable of extreme accuracy. Thus for example, it can incorporate a sophisticated contact tracing effort with multi-level two-way tracing, or predict the effects of a single unlicensed party. The disadvantages are enormous computational cost, lack of physical insight into the results and sensitivity of the outputs to the underlying network structures assumed by the modellers. Among prominent examples of agent-based models are References [1–3]. Another such model [32] has been able to explain the linear growth in corona cases seen in many regions of the world – linear disease trajectory is not observed in differential equation models except as a marginal case.
- Stochastic differential equation model: These attempt to combine the features of lumped parameter and agent based models, by writing differential equations which feature random variables [4, 7]. Our personal preference is for the preceding two kinds of models since they are more direct and intuitive.
- Data-driven model: These models simply take the existing data of COVID-19 spread over the past week or month (or longer) and use machine learning etc methods to generate a forecast for the next week or month. They pay little or no attention to the underlying processes driving the spread of the disease.

This concludes our summary of the different approaches in existence to the mathematical modeling of infectious diseases in general and COVID-19 in particular. What we propose in this article is a lumped parameter model; we shall now present a summary of the state of the art in this area.

1.2. Lumped parameter models

The first ever model for an infectious disease was of this type – it was the S-I-R model invented in 1927 by WILLIAM KERMACK and ALEXANDER MCKENDRICK [17] for generating the epidemiological curves of plague. A hundred years later, this model has been applied directly to COVID-19 as well. More common for this disease however is the S-E-I-R model, which we describe briefly. The four compartments here stand for susceptible, exposed, infected and removed (recovered or dead), and the equations are

$$\frac{dS}{dt} = -\frac{\beta_1 SI}{N}, \quad (1.1a)$$

$$\frac{dE}{dt} = \frac{\beta_1 SI}{N} - \beta_2 E, \quad (1.1b)$$

$$\frac{dI}{dt} = \beta_2 E - \beta_3 I, \quad (1.1c)$$

$$\frac{dR}{dt} = \beta_3 I, \quad (1.1d)$$

where the β 's are constants. Although all the β 's look like rates, they actually represent two kinds of real-world phenomena. β_1 for example is a true rate; it is the number of people whom one infected person infects in one unit of time. β_3 however is not a true rate but a reciprocal time – the fourth equation $dR/dt = \beta_3 I$ attempts to represent the fact that patients remain infected for an average duration of $1/\beta_3$ before being removed from consideration. This is a conceptual stretch; for if at $t = 0$ we have a certain number I_0 of infected people and set $\beta_2 = 0$ (by whatever means), then the subsequent infection profile looks like $I = I_0 \exp(-\beta_3 t)$. In reality however, if there are I_0 infected people today and no fresh infections, then they will all recover more or less uniformly within the next 10 or 14 days, or whatever is the infection period – there will not be a wave of recoveries at the start and a small but finite number of infected people several months down the line. This is a limitation of lumped parameter ODE models.

A second limitation of these models is that the parameters are often heuristic and not related directly to real-world phenomena governing the disease. For example, one may ask the question, “Suppose it turns out that the asymptomatic infection period, which was earlier thought to be 4 days, is actually 7 days, then which β 's would have to change and by how much?” or equivalently “Suppose the authorities of a university campus were to initiate mandatory weekly testing of all students, then which β 's would change and by how much?” These questions are all but impossible to answer for the basic S-E-I-R model (1.1).

More elaborate models with eight or more dependent variables and even more parameters [10, 15] have been set up in attempts to answer such questions – in such cases, the conceptual clarity of the model and the results arising therefrom potentially take a hit. For example, a recent study featuring such a model [6], led by an extremely distinguished mathematician, has been challenged [22] on the grounds that the parameters in the model “usually are not constrained in any [way] by our understanding of COVID-19 epidemiology” and that “a few of them are varied throughout a wide range, from 67 to 475000 in [one instance, with] no justification being given for this”. Whereas it is not our aim to critique either the original study or the rebuttal, we do believe that a model whose basic parameters are concrete things like asymptomatic fraction and pre-symptomatic infectious period is immune to attacks of this nature. Moreover, an ODE model with excessively many parameters can generate good fits to existing case trajectories at many distinct points or regions of parameter space – future trajectories with the different parameter sets can vary widely however. One study [31] which had made a rash prediction on such a flimsy basis was severely criticized [12] shortly after it appeared.

DDEs have been used in the Literature much more sparingly than ODEs. A notable example [35] dates from the pre-COVID era; this model has been followed up and applied to COVID by a different set of authors [33]. An elegant analytical solution to a linear DDE has been given in reference [11], which does not focus too much on the modeling aspects.

In this work we shall propose a lumped parameter model for the spread of an infectious disease (not necessarily just COVID-19) which has an elegant structure, is not over-ripe with parameters but at the same time directly incorporates the effects of realistic phenomena associated with the spread, such as pre-symptomatic transmission, asymptomatic carriers and test-trace-isolate programs. Our model further has the scope to accommodate different kinds of immune response to the disease. In Section 2 we present the derivation and solutions of the baseline model, in Section 3 we consider enhancements taking into account public health effects and in Section 4 we consider enhancements taking into account immunity effects. While much of the discussion has COVID-19 at its core, the features we model are quite general and are applicable to plague, influenza, Ebola virus disease, and any novel infectious disease which may emerge in future.

2. THE BASELINE MODEL

The treatment here closely follows our original works [26, 28, 29] in which we presented this model for the first time. We do not claim this section as a novelty of the present article, although we do believe that it is essential for what follows.

TABLE 1. List of parameters used in the baseline model.

| Symbol | Meaning |
|----------|--------------------------------------|
| m_0 | Per-case spreading rate |
| N | Initial susceptible population |
| μ_1 | Fraction of asymptomatic carriers |
| μ_3 | Fraction of cases not contact-traced |
| τ_1 | Asymptomatic infection period |
| τ_2 | Pre-symptomatic transmissible period |
| τ_h | $\tau_2/2$ |

2.1. Derivation

We present the derivation of the baseline model in more detail than just a summary because it will form the building block of all the advanced variants we shall present in Sections 3 and 4. We focus on a region with good mixing among its inhabitants, such as a neighbourhood, town, village or city (it might be advantageous to break up a big city into several regions though, depending on connectivity). Our model features only a single dependent variable $y(t)$ which is the cumulative count of corona cases in the region as a function of time. We measure time in days throughout this Article. Cases can be of two types – quarantined and at large. The former by definition have zero contribution to spread. The latter disseminate the virus freely to healthy and susceptible people *via* interaction. By interaction we mean both in-person interaction and interaction *via* objects (for example, a case contaminating an item at a grocery store which a subsequent healthy customer buys qualifies as an interaction). In what follows, we shall use the word *target* to denote a person with whom an at large case interacts. The spread of the disease *via* interaction can be represented through the word equation

$$\begin{aligned} \text{(Rate of emergence of new cases)} &= \text{(Spreading rate of each at large case)} \\ &\times \text{(Probability of target susceptibility)} \times \text{(Number of at large cases)}, \end{aligned} \quad (2.1)$$

which we shall now express in a mathematical form. Before that, we enumerate in Table 1 the parameters which shall be of relevance, so that a reader may skip the details of the derivation and proceed to the final equation (2.3) if s/he so desires.

Every at large case has a certain rate of interaction with targets. This interaction rate can vary widely – for example a grocer or banker might deal with 15–20 customers a day while a working from home software engineer might not interact with anyone at all in a week. For a lumped parameter model, we must average over various kinds of people; let the average interaction rate over all at large cases be q_0 persons/day. This q_0 depends on the degree of mobility in society – in a lockdown it will be far lower than in an unlocked state. Not every interaction however results in a transmission – for example a target sitting in the same classroom as a case might not contract the disease if both are wearing masks and sitting far enough apart. Similarly, a contaminated grocery item will not spread the virus to the customer if s/he sanitizes everything s/he buys. Let P_0 be the probability (averaged over cases and targets) that an interaction between a case and a susceptible target actually results in a transmission. Note again the definition – *given that the target is susceptible*, P_0 denotes the probability that s/he contracts the virus from the case. Then, we define the per-case spreading rate $m_0 = q_0 P_0$. We can see that m_0 incorporates the effects of public health interventions.

Since P_0 is conditional on the target’s being susceptible, we now need to account for the possibility that s/he is not. The second term on the RHS of (2.1) factors this in. For the baseline model we assume permanent immunity *i.e.* we assume that recovered cases are insusceptible to further infection for all time. With permanent immunity, the probability of a random person’s being insusceptible is the total number of recovered cases plus the total number of at large cases divided by the total number of people in the region. At large plus recovered cases add up to approximately y – the count is actually a little less than y since not all cases at any time are

either at large or recovered (the exceptions are the active cases in quarantine). But if the quarantine duration is significantly less than the overall progression of the disease, then we can make the *zero quarantine* assumption and treat the recovery count as y . Let the total number of susceptible people in the region at the start of the epidemic be N . Then the probability of a random person's being insusceptible is y/N and the probability of his/her being susceptible is $1 - y/N$. The zero quarantine assumption ensures that our model possesses the very important property of dy/dt being identically zero when $y = N$, without having to incorporate a lot of fancy end-effects.

To motivate the third term, suppose hypothetically that all cases take a time τ days to recover and remain at large the whole time. Then the number of at large active cases now is exactly the number of people who fell sick between now and τ days back – mathematically, this is expressed as $y(t) - y(t - \tau)$. Here is the delay in the equation – note that we are using *delay* to express the recovery duration rather than an inverse rate, as in the S-E-I-R and related models. Of course, every case having the same recovery time and remaining at large throughout is simplistic. For one, symptomatic cases (with a few negligent exceptions) will generally go into quarantine after manifesting symptoms. For another, contact tracing drive will yield and isolate asymptomatic as well as pre-symptomatic or pre-symptomatic cases. We now partition the cases into three classes : (a) contact traced cases, (b) untraced symptomatic cases and (c) untraced asymptomatic cases. Let the number μ_1 between 0 and 1 denote the fraction of total cases who are asymptomatic, let μ_3 between 0 and 1 denote the fraction of total cases who do not get contact traced, let τ_1 denote the asymptomatic infection period and τ_2 the pre-symptomaticity period during which a to-be-symptomatic case is transmissible prior to developing symptoms.

The class (a) or contact traced cases account for $1 - \mu_3$ of the total. This is a quasi-heuristic rather than a phenomena-driven parameter which suffices for the baseline model. For this model, we also make two assumptions whose effects counteract each other – the first is the assumption of *zero non-transmissible incubation period (NTI)* and the second is that of *instantaneous contact tracing*. The first assumption implies that if a target contracts the virus, then s/he begins transmitting immediately following exposure. The second assumption implies that public health authorities track down a person's contacts as soon as they begin the tracing process. We have used these assumptions so as not to clutter the baseline model with parameters – in Section 3.2 we shall show how to relax them. If the contact tracing starts from freshly reporting symptomatic cases, then, with the assumptions in place and taking for granted that all cases transmit continuously and uniformly, the average duration that the secondary cases of the reporting cases remain at large is $\tau_2/2$. This happens because the tracing picks up cases who were exposed by the symptomatic case a time τ_2 ago (when that case just turned symptomatic) as well as cases who were exposed very recently, and everyone in between. Class (b) or untraced symptomatic cases account for fraction $\mu_3(1 - \mu_1)$ of the total cases and they remain at large for the pre-symptomatic period τ_2 before manifesting symptoms and going into quarantine (the baseline model does not account for the actions of negligent individuals who we hope are few in number). Finally, class (c) or untraced asymptomatic cases account for fraction $\mu_3\mu_1$ of the total cases and these remain at large for the asymptomatic infection period τ_1 . Applying the $y(t) - y(t - \tau)$ argument to each class yields the mathematical representation of the third term on the RHS of (2.1) as where we use τ_h to denote $\tau_2/2$,

$$n = (1 - \mu_3) (y - y(t - \tau_h)) + (1 - \mu_1)\mu_3(y - y(t - \tau_2)) + \mu_1\mu_3(y - y(t - \tau_1)). \quad (2.2)$$

Multiplying all the terms on the RHS of (2.1) and simplifying the algebra in (2.2), we get

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y}{N} \right] [y(t) - (1 - \mu_3) y(t - \tau_h) - (1 - \mu_1) \mu_3 y(t - \tau_2) - \mu_1 \mu_3 y(t - \tau_1)], \quad (2.3)$$

which is the *retarded logistic equation*, and the baseline model.

2.2. Solutions

We first note that (2.3) enables a direct calculation of the reproduction number R at any stage of evolution of the disease. R represents the number of people to whom one person transmits the disease, on the average. It

is a very important parameter characterizing disease spread since the epidemic grows if $R > 1$ and dies down if $R < 1$. The formula we find is

$$R = m_0 \left(1 - \frac{y}{N}\right) \left(\frac{1 + \mu_3 - 2\mu_1\mu_3}{2}\tau_2 + \mu_1\mu_3\tau_1\right). \quad (2.4)$$

The derivation of this has been reported in our prior work [29, 30] and is given in Appendix A. We note that the calculation of R from an ODE model [13] is non-trivial, and is generally limited to the starting value R_0 only.

We now briefly present six different classes of solutions of (2.3), which are manifest for different choices of parameter values. These choices are not derived from fits to specific data sets but rather are chosen to demonstrate the different kinds of behaviour which the epidemic trajectories can show with this model. Extensive fits of real country data to the model have been demonstrated in Reference [19]. We use numerical integration to solve (2.3), the method being second order Runge Kutta with a step size of 0.001 day. While we have no idea regarding the accuracy of the solver for the present equation, we have calibrated it by solving equations with known analytical solutions and have found errors of 0.1 percent or less. We solve (2.3) in a Notional City having an initial susceptible population of $N = 300000$ and an initial condition of zero cases to start with and 100 cases/day for the first seven days; such a City will be our solution domain in the entire Article. For all Cities in this Section, we assume $\mu_1 = 0.8$. The parameter values of the six Cities A to F as well as real-world examples of each city are given in Table 2, while the solutions for these Cities are shown in Figure 1.

In each plot of the below Figure, we show three things – the case count $y(t)$ as a blue line, its derivative $\dot{y}(t)$ as a green line and the weekly case growth scaled down by a factor of 7 in grey bars. This latter is called the epidemiological curve or “epi-curve”.

This shows us that the baseline model is itself capable of generating a diverse range of epidemic trajectories which we see in the real world. Note that the various parameters can be estimated either from medical information about the disease [8] or from empirical experiments [23].

3. PUBLIC HEALTH INTERVENTION EFFECTS

Here we build upon the baseline model to accommodate the effects of various public health interventions. Before beginning the generalizations, we note three features of the baseline equation (2.3) which represents the word equation (2.1). The first is that it represents a *clean decoupling* of the various classes of effects – public health terms come into the first term on the RHS, immune response effects into the second term and intrinsic disease characteristics such as asymptomatic fraction and symptomatic pre-symptomatic period into the third term. This is not the case for the S-E-I-R model. This feature implies that the question we posed in Subsection 0.2 about the effects of changing the asymptomatic infection period from 4 days to 7 days can be answered in one line basis of (2.3) [our second question is more involved and the answer is coming later in this Section].

The second noteworthy feature of the baseline equation is that it is cast in terms of a *single variable* rather than an array of variables. Quantities such as the numbers of hospitalizations and deaths can be easily extracted from the cumulative case trajectory by using the known hospitalization and mortality rates and intervals between contraction and recovery or death. Thus, the smaller number of variables does not amount to a limitation on the information that can be obtained from our model. Rather, by using DDE instead of ODE, it is possible to create a versatile model with a less complex appearance.

The third feature of (2.3) is its *negligible computational cost*. The code for solving (2.3) takes about 100 lines of Matlab to write, and about one second on a laptop to run. The sizes and runtimes will remain of the same order for all the variations we shall consider in this Article. With this, we go on to our consideration of the individual public health intervention steps.

TABLE 2. Six solution classes of the baseline model (2.3). These solutions have been chosen to demonstrate different scenarios such as free evolution of the disease under different degrees of social mobility (Cities A through C), effect of a late lockdown (City D), reopening without increase in cases (City E) and reopening with second wave (City F). These solutions correspond to different kinds of COVID-19 trajectories seen all over the world. The values of m_0 and μ_3 in each case have been chosen somewhat arbitrarily, to generate a desired temporal profile of the reproduction number R . This is on account of our observation that to the largest order the disease dynamics is governed by R rather than its constituent details. We would like to highlight that the solution classes in this Table and in Figure 1 serve primarily to demonstrate the versatility of our model rather than provide detailed fits for policymakers etc. The latter exercise has been conducted by Kreck and Scholz [19].

| City | m_0 | μ_3 | R as a function of time | Description | Example |
|------|---|-------------------------------------|--|--|--|
| A | 0.23 | 0.5 | Less than unity throughout | Epidemic contained <i>via</i> public health measures | All cities in New Zealand |
| B | -do- | 0.75 | Starts at 1.15 and evolves naturally | Epidemic grows initially, reaching partial herd immunity [20] | Most or many cities in India and in California (USA) |
| C | 0.5 | -do- | Starts at 2.5 and evolves naturally | Epidemic progresses to herd immunity with little intervention attempted | None |
| D | 0.5 upto 40,000 cases and 0.23 thereafter | -do- | Starts at 2.5 but reduces to below unity <i>via</i> lockdown | Late lockdown causes initial explosion before settling down to partial herd immunity | New York City, USA |
| E | 0.23 upto 80 days and 0.5 thereafter | 0.5 upto 80 days and 0.1 thereafter | Remains less than unity throughout | Advanced contact tracing drive enables significant reopening with no spreading | None |
| F | -do- | 0.5 upto 80 days and 0.2 thereafter | Becomes greater than unity on the 80th day and evolves naturally | Reopening causes a second wave | Cities in UK, Germany |

3.1. Age, vulnerability and transmissibility structuring

Our first generalization deals with structuring. The most common structuring is age-structuring which is important because the effects of COVID-19 on young and old people are vastly disparate (the latter are much more severely). Let y_1 denote the cumulative case count in the young population and y_2 the case count in the old population. Instead of a single interaction rate q_0 , we now need three interaction rates : q_1 , that of young people with other young people, q_2 , that of young people with old people and q_3 , that of old people with old people. Contact matrices have been reported in Literature [23]. Let N_1 and N_2 denote the initial susceptible numbers of young and old people. Writing (2.1) for the young people and expressing it mathematically as in (2.3), we get

$$\begin{aligned} \frac{dy_1}{dt} = & q_1 P_0 \left[1 - \frac{y_1}{N_1} \right] [y_1(t) - (1 - \mu_3) y_1(t - \tau_h) - (1 - \mu_1) \mu_3 y_1(t - \tau_2) - \mu_1 \mu_3 y_1(t - \tau_1)] \\ & + q_2 P_0 \left[1 - \frac{y_1}{N_1} \right] [y_2(t) - (1 - \mu_3) y_2(t - \tau_h) - (1 - \mu_1) \mu_3 y_2(t - \tau_2) - \mu_1 \mu_3 y_2(t - \tau_1)]. \end{aligned} \quad (3.1)$$

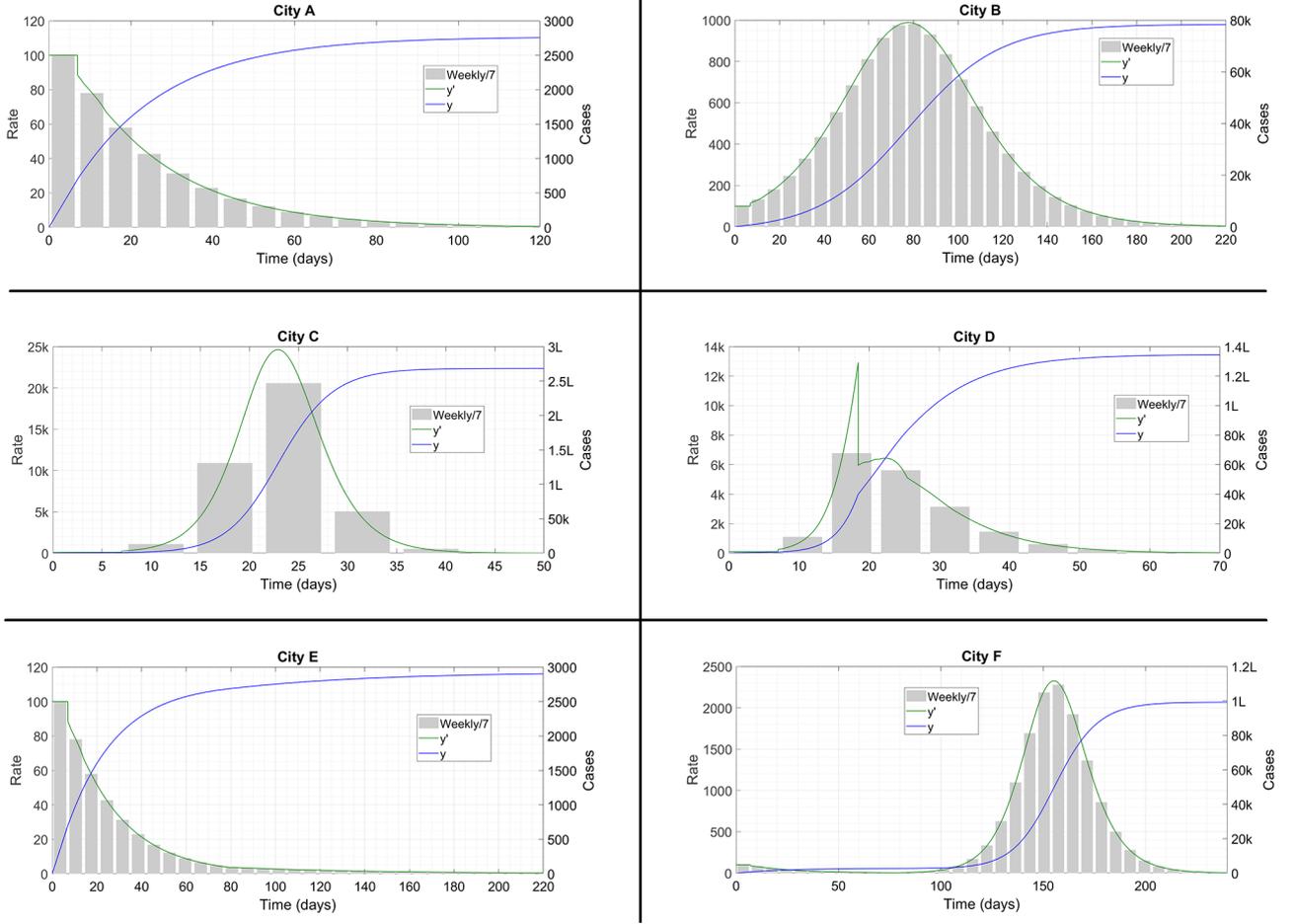


FIGURE 1. Case curves for the solution classes mentioned in Table 1. Adapted from Reference [29]. The symbol ‘k’ denotes thousand and ‘L’ hundred thousand.

Recall that P_0 is the probability that an interaction between a case and a susceptible target actually results in a transmission. The first line on the above RHS denotes the transmission to young targets from young cases while the second line denotes transmission to young targets from old cases. Note that the susceptibility probability is $1 - y_1/N_1$ in both terms – since q_1 and q_2 already accommodate the fact that the target is a young person, the thing to ask here is, given that the target is young, what is the probability that s/he is not immune. By analogy, the case counts among the old people will be given by

$$\begin{aligned} \frac{dy_2}{dt} = & q_2 P_0 \left[1 - \frac{y_2}{N_2} \right] [y_1(t) - (1 - \mu_3) y_1(t - \tau_h) - (1 - \mu_1) \mu_3 y_1(t - \tau_2) - \mu_1 \mu_3 y_1(t - \tau_1)] \\ & + q_3 P_0 \left[1 - \frac{y_2}{N_2} \right] [y_2(t) - (1 - \mu_3) y_2(t - \tau_h) - (1 - \mu_1) \mu_3 y_2(t - \tau_2) - \mu_1 \mu_3 y_2(t - \tau_1)]. \end{aligned} \quad (3.2)$$

This can be generalized to include many more vulnerability classes. Most existing ODE disease models can be extended for age structuring however (the logistic model is a notable exception), so this is not a significant novelty of our DDE model.

We now consider transmissibility structuring. Suppose that among the N interacting people, N_1 wear masks all the time and N_2 don't wear them at all ($N_1 + N_2 = N$). This kind of situation is very common especially in USA where there is considerable resistance to mask-wearing in some places. A structured model is superior to a univariate model here for two reasons : (a) it can enable public health authorities to quantitatively determine how much better the mask wearers will fare and accordingly formulate public policy, and (b) it is possible that mask wearers might suffer less severe symptoms due to receiving smaller viral loads, so the hospitalization and mortality characteristics might be different among the two groups (whether this is true or not for COVID-19 is currently unknown – a recent study [5] reports a small but statistically significant correlation between mask mandates and hospitalization rates, and we haven't seen anything more definitive).

The raw input for this situation is a set of four probabilities : the probability P_1 that the disease jumps from a masked case to a masked target when they interact, the probability P_2 that the disease jumps from a masked case to an unmasked target, the probability P_3 that the disease jumps from an unmasked case to a masked target and the probability P_4 that the disease jumps from an unmasked case to an unmasked target. Common sense says that P_1 will be the smallest and P_4 the greatest of the four probabilities; studies seem to indicate that P_2 is less than P_3 *i.e.* one mask between two people is more useful when on the case than on the target. We assume that the interaction rate is constant *i.e.* every person interacts at the same rate q_0 with other persons, whether masked or otherwise. We also ignore age-structuring. This last statement is a general philosophy we shall adopt throughout this Article – *when considering each new situation, we shall incorporate the modification directly into the baseline model and not a variant form.* Thus, we shall not derive an equation with both age and mask structuring.

To model this situation, let y_1 denote the cumulative number of cases in the masked population and y_2 the same thing in the unmasked population. Considering the y_1 -dynamics first, there will be two terms just as in (3.1) to account for transfers to masked targets from masked and unmasked cases. For each term, we again have the structure (2.1). By our assumptions, masked cases interact with all others at an average rate q_0 , and since all interactions take place equally, a fraction N_1/N of these will be with masked targets. Then, we have the probability P_1 of mask to mask transmission, the probability $1 - y_1/N_1$ of masked target susceptibility and a set of delay terms for the number of masked cases at large. For the second term, unmasked cases again interact with everyone at rate q_0 and fraction N_1/N of these interactions are with masked people. Then we have the probability P_3 of transmission from unmasked case to masked target, followed by the susceptibility probability as above and a second set of delay terms counting the unmasked cases at large. This implies

$$\begin{aligned} \frac{dy_1}{dt} = & \frac{q_0 N_1 P_1}{N} \left[1 - \frac{y_1}{N_1} \right] [y_1(t) - (1 - \mu_3) y_1(t - \tau_h) - (1 - \mu_1) \mu_3 y_1(t - \tau_2) - \mu_1 \mu_3 y_1(t - \tau_1)] \\ & + \frac{q_0 N_1 P_3}{N} \left[1 - \frac{y_1}{N_1} \right] [y_2(t) - (1 - \mu_3) y_2(t - \tau_h) - (1 - \mu_1) \mu_3 y_2(t - \tau_2) - \mu_1 \mu_3 y_2(t - \tau_1)], \end{aligned} \quad (3.3)$$

and by analogy

$$\begin{aligned} \frac{dy_2}{dt} = & \frac{q_0 N_2 P_2}{N} \left[1 - \frac{y_2}{N_2} \right] [y_1(t) - (1 - \mu_3) y_1(t - \tau_h) - (1 - \mu_1) \mu_3 y_1(t - \tau_2) - \mu_1 \mu_3 y_1(t - \tau_1)] \\ & + \frac{q_0 N_2 P_4}{N} \left[1 - \frac{y_2}{N_2} \right] [y_2(t) - (1 - \mu_3) y_2(t - \tau_h) - (1 - \mu_1) \mu_3 y_2(t - \tau_2) - \mu_1 \mu_3 y_2(t - \tau_1)], \end{aligned} \quad (3.4)$$

which complete the formulation of the transmissibility-structured model.

In Figure 2 we show a sample plot of the solutions of (3.3) and (3.4) for the parameters $\tau_1 = 7$ and $\tau_2 = 3$ (as for all simulations in this Article), $\mu_1 = 0.8$, $\mu_3 = 0.75$, $q_0 = 4$, $P_1 = 1/30$, $P_2 = 1/20$, $P_3 = 1/15$, $P_4 = 1/3$, $N_1 = 250000$ and $N_2 = 50000$. In these and similar plots, we use the same colours for y_1 as used for y in Figure 1, and red for cases, magenta for derivative and cyan bars for epi-curve of y_2 . We can see about 55,000 cases among the masked population which amount to just above 20 percent of the total 2,50,000 masked people. By

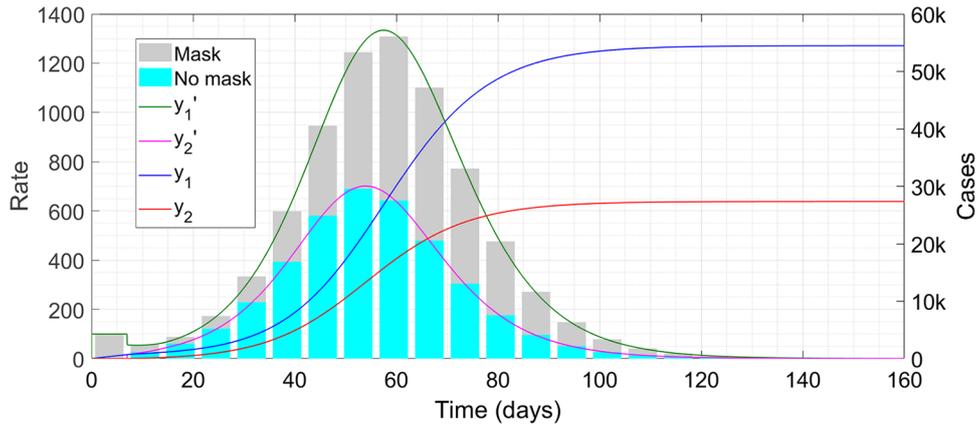


FIGURE 2. Case trajectories with extensive mask use. The symbol ‘k’ denotes thousand.

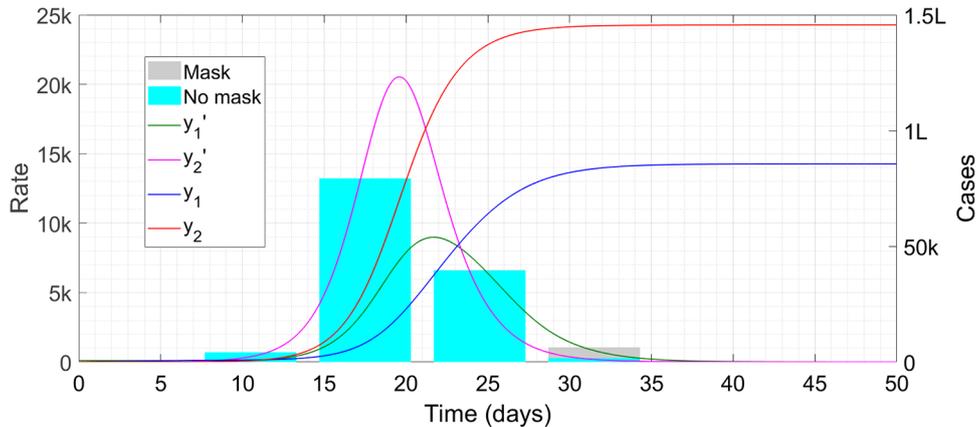


FIGURE 3. Case trajectories with 50 percent mask use. The symbol ‘k’ denotes thousand and ‘L’ hundred thousand. Note that the cyan bars obscure the grey bars in many places; the envelope of the latter is the green curve which is still visible.

contrast there are about 27,000 unmasked cases which amount to more than 50 percent of the 50000 unmasked people.

We now run the code again, this time setting $N_1 = N_2 = 150000$ *i.e.* equal numbers of masked and unmasked people (Fig. 3).

Not only is the unmasked population infected almost entirely but also there is more than 50 percent infection level among the masked people. This type of statistic can be used by public health authorities to encourage mask use – by not masking, not only are you increasing your own chances of catching corona but you are subjecting the law-abiding people to extra risk as well. Of course, for a propaganda drive, we first need good-quality data on the efficacy of mask use – the values of P_1 to P_4 we used here for the simulations were, though plausible, arbitrary. Two sources which can yield these values are (a) laboratory experiments and simulation studies, and (b) follow-ups of exposures arising from contract tracing activities, including cases as well as non-cases who were exposed to a known transmitting case.

3.2. Contact tracing

This is one of the more difficult aspects to incorporate into a lumped parameter model, and is an area where agent-based models have an intrinsic advantage. The DDE however is capable of taking this in its stride. As we have already seen, the baseline model itself incorporates contact tracing to a fair extent. It contains two assumptions however – those of zero non-transmissible incubation period (NTI) and instantaneous tracing – which cannot be included in a more detailed model where contact tracing is a priority. It also contains a heuristic parameter μ_3 which is too coarse for the present application.

The model derivation in this case is quite involved so we present it in Appendix B. Here we proceed straight to the final equation. We assume one-level forward contact tracing starting from freshly reporting symptomatic cases. The parameters involved are

- The non-transmissible incubation (NTI) period τ_4 .
- The probability X^* that a random case is infected by a symptomatic case.
- The probability P_3 that the contacts of a known symptomatic case are actually picked up during tracing.
- The time τ_3 it takes to trace a case's contacts.

These go into determining the time T for which a random contact traced case remains at large before being caught and quarantined. The final equation turns out to be

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y(t - 2\tau_4)}{N} \right] \times [y(t - \tau_4) - P_3 X^* y(t - T - \tau_4) - (1 - \mu_1)(1 - P_3 X^*) y(t - \tau_2 - \tau_4) - \mu_1 (1 - P_3 X^*) y(t - \tau_1 - \tau_4)], \quad (3.5)$$

where T and X^* have complicated mathematical expressions which are given in the Appendix.

Since (3.5) has nearly the same form as the baseline, we refrain from showing simulation time traces. Rather, we give the values of T , X and hence the capture probability $P_3 X$ or $1 - \mu_3$ for some realistic sets of parameter values. With the standard values 7 and 3 for τ_1 and τ_2 and taking $\tau_4 = 4$ (the total incubation period is approximately 7 days on average [8]), the value $\tau_3 = 2$ gives $T = 1/6$. This implies that contact traced cases remain at large for a relatively short time. The fraction of cases which is caught depends highly on two parameters – the asymptomatic fraction μ_1 and the tracing success probability P_3 . Thus, if 30 percent cases are asymptomatic, then $P_3 = 1/2$ leads to capture of 25 percent of the total cases while $P_3 = 9/10$ leads to capture of 46 percent of the total cases. If 80 percent cases are asymptomatic however, then $P_3 = 1/2$ captures not even 5 percent of the total cases, which increases to 8.7 percent if P_3 is raised to $9/10$.

This is not a surprise. Since the contact tracing is being carried out starting from symptomatic cases, the process efficiency automatically becomes crippled if they are few in number. Tracing tertiary contacts *i.e.* contacts of secondary cases will not really help since the secondary cases get quite a low air-time anyway. Two-way contact tracing, *i.e.* when a new case reports symptomatic then trying to identify the case's source of exposure, and again following up on that source's other exposures can lead to the capture of many more cases. This process is difficult however since an exposure typically occurs 7 days prior to symptoms, and a typical patient will hardly be expected to remember his/her movements from 7 days past. Another way of capturing asymptomatic cases is by performing regular mass testing; this we discuss below.

3.3. Mass testing

Periodic testing of the entire population is a disease mitigation strategy which is popular in organizations with small staffing and large funding. To model this effect, we start as usual from the underlying structure (2.1). As with contact tracing in the last Section, the first two terms remain unchanged – what frequent testing changes is the duration for which cases remain at large. To lend concreteness to the discussion, we introduce numerical values at this step itself. Let each person be tested once a week, let the test itself be perfectly sensitive and let the results take one day to come through. The asymptomatic and pre-symptomatic periods remain $\tau_1 = 7$

and $\tau_2 = 3$ as they always were. Let Mr X be an asymptomatic case who does not get contact traced and has his test on a Monday. With equal probability $1/7$, Mr X can turn transmissible on any day of the week. If he starts transmitting on a Monday, he transmits for only one day before he is detected and quarantined. If he starts on a Sunday, he goes 2 days at large before capture, if Saturday then 3 days and so on until if he turns transmissible on a Tuesday, bad luck, he spends the entire 7-day asymptomatic infection period at large. The average duration for which he remains at large is $(1/7)(1 + 2 + 3 + 4 + 5 + 6 + 7)$ which is 4 days. Similarly let Ms Y be an untraced symptomatic case who gets tested on Monday. She transmits for 1 day if she starts on a Monday, 2 days if Sunday and the full pre-symptomaticity period of 3 days otherwise. Thus, the average duration for which she remains at large is $(1/7)(1 + 2 + 3 \times 5)$ which is $18/7$ days.

For contact tracing, we go back to the simple structure of the baseline model instead of using the complications of the last Section – a heuristic parameter μ_3 for the escape fraction and the assumptions of zero NTI and instantaneous tracing. This time however, tracing will proceed from every case which is picked up during testing. The traced cases will spend an average at large time of half their sources; these latter spend $4\mu_1 + (18/7)(1 - \mu_1)$ days. Thus we have the model

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y}{N} \right] [y(t) - (1 - \mu_3)y(t - T) - (1 - \mu_1)\mu_3y(t - 18/7) - \mu_1\mu_3y(t - 4)], \text{ where} \quad (3.6a)$$

$$T = 2\mu_1 + \frac{9}{7}(1 - \mu_1). \quad (3.6b)$$

Once again, we have the same basic structure as (2.3) with new parameter values. This answers the second question with which we had challenged the S-E-I-R model in Subsection 0.2. In Appendix C we show how to extend this calculation to the case where the test has imperfect sensitivity.

Since the equation (3.6a) is similar in structure to (2.3), we again refrain from showing simulations. This time, a relevant question is, what is the effect of the testing drive on the reproduction number R . Repeating the procedure which led to (2.4) shows that R for (3.6a) is given by

$$R = (1 - \mu_3)T + (1 - \mu_1)\mu_3T_2 + \mu_1\mu_3T_1 \quad (3.7)$$

Using the standard values of τ_1 and τ_2 , $\mu_1 = 0.8$ and $\mu_3 = 0.9$, we calculate the reduction in R if the weekly testing drive is implemented while keeping everything else the same. We find a 38 percent reduction if the test is 100 percent sensitive and a 22 percent reduction if the test is 60 percent sensitive. If $\mu_1 = 0.3$ and $\mu_3 = 0.5$, then the reduction drops to 21 percent with fully sensitive test and 8.5 percent with 60 percent sensitive test. The decrease from the previous situation is not surprising since the weekly testing does not significantly reduce the at large duration of symptomatic cases. The value of $\mu_1 = 0.3$ is too low however and the earlier set of figures is more realistic. A 20-30 percent reduction in R is not jaw-dropping but not trivial either – it is ineffective against a really high R of say 2 but it can bring down a somewhat-above-unity R to just below unity and extinguish the epidemic in time.

Again we note that although the interventions of masking, contact tracing and testing are all currently relevant in the context of corona, they are applicable to other infectious diseases as well as such as pandemic influenza and will very likely be used whenever a new infectious epidemic breaks out anywhere in the world – north or south, west or east.

4. IMMUNE RESPONSE EFFECTS

While the last Section dealt with public health interventions, in this Section we turn to another aspect of the disease which is the human immune response. The baseline assumption is that of permanent immunity – one bout of infection renders a person insusceptible to the disease for all time. This assumption is valid whenever the immunity period is longer than the overall progression of the epidemic – immunity need not be genuinely

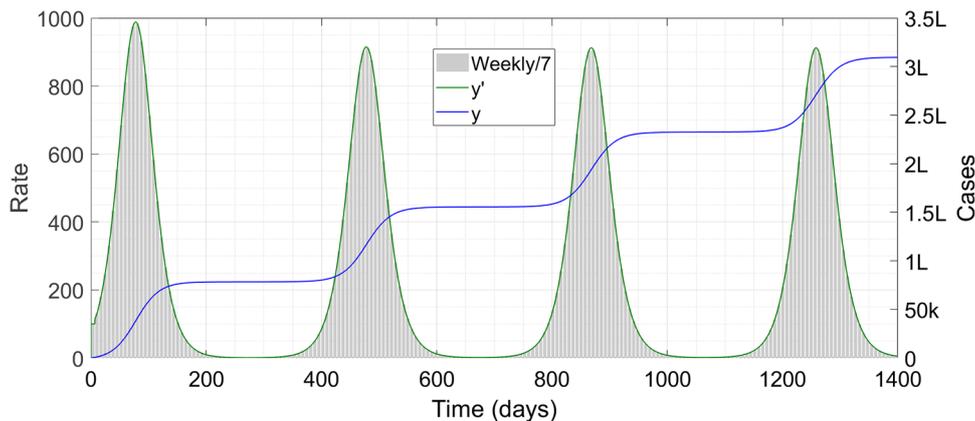


FIGURE 4. Case trajectories with temporary immunity. The symbol ‘k’ denotes thousand and ‘L’ hundred thousand.

lifelong. The question naturally arises as to what profile the spreading trajectories may take if immunity really turns out to be temporary in a significant fraction of the population.

4.1. Simple temporary immunity

Here we consider the case where immunity against the disease lasts for a fixed, limited duration following the first infection. After this duration is over, the person becomes susceptible to the disease again. One study [14] finds such a situation for benign human coronaviruses (not SARS, MERS or COVID-19). Let τ_0 be the duration for which immunity lasts. As is customary in lumped parameter models, the value we use must be the average over the entire population.

As usual, we start from the word equation (2.1). Here, the second term on the RHS denotes the immune response so only that will be altered relative to the baseline model (2.3). The term must still denote the probability that a random person is susceptible. By our assumptions, each recovered case remains immune or insusceptible for a duration τ_0 , the immunity cutoff period, typically of the order of weeks or months. So, if a person once contracts and recovers from the infection at time t , then s/he remains immune upto time $t + \tau_0$ and then becomes susceptible again. At any given instant, the people who are immune are all those who have contracted the infection during the last τ_0 days, and no one else. Hence, the number of insusceptibles at time t is exactly the number of new infections which have occurred between time $t - \tau_0$ and t , which is $y(t) - y(t - \tau_0)$. The structure of this term and the underlying logic are the same as those motivating the other delay terms in (2.3). In proposing this structure, we have assumed that recoveries are instantaneous (compared to τ_0 , also see Sect. 2.1) and have also ignored deaths. Since the mortality rate of COVID-19 is fortunately quite low, this second assumption is reasonable as well. Then, the probability that a random person at time t is susceptible is

$$P = 1 - \frac{y(t) - y(t - \tau_0)}{N}, \quad (4.1)$$

and working this into (2.1) yields

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y(t) - y(t - \tau_0)}{N} \right] [y(t) - (1 - \mu_3)y(t - \tau_h) - (1 - \mu_1)\mu_3y(t - \tau_2) - \mu_1\mu_3y(t - \tau_1)]. \quad (4.2)$$

This is the generalization of the baseline model to the case where immunity is temporary.

We run a simulation with parameters $m_0 = 0.23$, $\mu_1 = 0.8$, $\mu_3 = 0.75$ and $\tau_0 = 200$ days. The result is shown in Figure 4. We can see multiple waves of disease in this case, which is consistent with what has been found in

the prior literature [18, 24] using different models. Note that the waves continue indefinitely; we have artificially stopped the simulation at $t = 1400$ days.

4.2. Complex immune response

The assumption that every case becomes completely susceptible after a certain time is perhaps over-simplified. Specially for COVID-19, we do not have nearly enough data about the immune response to the pathogen. However, for diseases in general, there are three kinds of immunity – *sterilizing* immunity which completely prevents reinfection, *severity-reducing* immunity which mitigates the symptoms during reinfection and *transmissibility-reducing* immunity which mitigates the patient’s transmissibility during reinfection. Studies on the benign coronaviruses [21] have found that although sterilizing immunity wanes in about a year, severity-reducing immunity lasts much longer. In addition to the above, there is a dangerous phenomenon called *antibody-dependent enhancement (ADE)* in which case a reinfection takes a more severe form than the original infection. ADE has not been found for benign coronaviruses; it is unknown whether it occurs with COVID-19.

We consider a situation where every recovered case has the probability P_1 of remaining immune for life, P_2 of remaining fully immune for τ_0 days and then becoming susceptible to a lower virulence form (LVF) of the disease and P_3 (with $P_1 + P_2 + P_3 = 1$) of remaining fully immune for τ_0 days and then becoming susceptible to a higher virulence form (HVF) of the disease through ADE. We further assume that after catching the infection twice, a person does not contract it any longer (becomes either immunized or dead). To model this situation, we define three variables y , z_1 and z_2 . Here y denotes the count of cases in the current form of the disease, z_1 denotes the count of LVF cases and z_2 the count of HVF cases.

We start from the y -equation. Any person who catches the current form of the disease once is insusceptible to it for all future time – s/he contracts either nothing, or LVF, or HVF. Thus the susceptibility probability in this equation will be $1 - y/N$ as in the baseline model. As in the coupled models of Section 3.1, the number of at large cases will include all the forms of the disease, so there will be three sets of delay terms. We assume, for the sake of notational elegance more than anything else, that LVF and HVF have the same infection parameters as the current form. We then have

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y}{N} \right] n(y, z_1, z_2), \text{ where} \quad (4.3a)$$

$$\begin{aligned} n(y, z_1, z_2) = & y(t) - (1 - \mu_3) y(t - \tau_h) - (1 - \mu_1) \mu_3 y(t - \tau_2) - \mu_1 \mu_3 y(t - \tau_1) \\ & + z_1(t) - (1 - \mu_3) z_1(t - \tau_h) - (1 - \mu_1) \mu_3 z_1(t - \tau_2) - \mu_1 \mu_3 z_1(t - \tau_1) \\ & + z_2(t) - (1 - \mu_3) z_2(t - \tau_h) - (1 - \mu_1) \mu_3 z_2(t - \tau_2) - \mu_1 \mu_3 z_2(t - \tau_1). \end{aligned} \quad (4.3b)$$

Here, the notation $n(y, z_1, z_2)$ is just a shorthand for a cumbersome but transparent expression.

For the z_1 -equation, we ask the usual question that what is the probability of a random person’s catching the LVF disease. For this we have to first count the number of people who are eligible to catch the LVF. This includes everyone who has already contracted the default form, and that a time τ_0 or more days ago, which is $y(t - \tau_0)$. From this eligible pool however, we shall have to exclude all those who have already had the second infection, either LVF or HVF. So the pool of eligible targets has the size $y(t - \tau_0) - z_1(t) - z_2(t)$. Given that a person is eligible, the probability of his/her catching LVF is P_2 . Putting all this together, the probability of a random person’s being susceptible to LVF is

$$P = P_2 \frac{y(t - \tau_0) - z_1 - z_2}{N}, \quad (4.4)$$

which leads to the equation for z_1 as

$$\frac{dz_1}{dt} = m_0 P_2 \left[\frac{y(t - \tau_0) - z_1 - z_2}{N} \right] n(y, z_1, z_2). \quad (4.5)$$

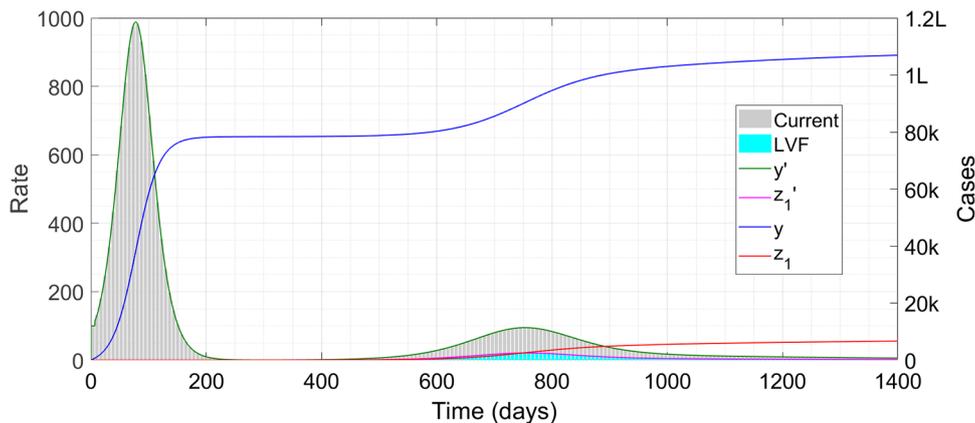


FIGURE 5. Case trajectories with given probabilities of different kinds of immune response. The symbol ‘k’ denotes thousand and ‘L’ hundred thousand.

It is now a simple matter to construct the equation for z_2 , where everything remains the same as above except that the probability P_2 is replaced by P_3 , giving

$$\frac{dz_2}{dt} = m_0 P_3 \left[\frac{y(t - \tau_0) - z_1 - z_2}{N} \right] n(y, z_1, z_2). \quad (4.6)$$

Equations (4.3a,4.5,4.6) constitute a coupled set of DDEs for the dynamics of the current and the modified virulence forms.

For the simulation run, we consider the same parameters as in Section 4.1 above and take $P_1 = 2/5$, $P_2 = 3/5$ and $P_3 = 0$. Thus, every person who has had one bout is either permanently insusceptible (40 percent chance) or susceptible to the LVF after 200 days (60 percent chance). The run is below, in Figure 5. We can see that instead of multiple identical waves as in Figure 4, there is only a very broad and low second wave. The first wave is entirely in the current form of the disease while the second features both the current and the LVF forms. Because the probability of catching LVF is quite low, the wave in that form is even smaller than in the current form. As in Figure 4, the runs continue beyond 1400 days but we have stopped it at that time. The total numbers of y and z_1 are bounded by N and $3N/5$ *i.e.* 300000 and 180000 respectively, which automatically prevents infinite perpetuation of the waves; the wave shapes are modified to fit this constraint.

Most of the assumptions we have made in this Section can be relaxed very easily. Incorporating different transmissibilities and durations for the three forms of infection is trivial, while working round the maximum two infections constraint is also easy. For the latter situation, we introduce more forms of the disease, say ultra-LVF w_1 and ultra-HVF w_2 and calculate the susceptibility probability of each person to the various infections to use in the structure (2.1).

5. LIMITATIONS, CONCLUSION AND FUTURE DIRECTIONS

Lumped parameter models do have certain intrinsic limitations. While ours is particularly versatile, there are some places where it still runs up against a wall. For example, the parameters q_0 , P_0 , τ_2 etc. in the baseline model have to be averaged over all at large cases. During contact tracing we were forced to make assumptions regarding continuous and uniform transmission etc. For these reasons, DDE cannot replace agent-based models. There are certain questions which a lumped parameter model can never answer. “Today there are exactly 5 active cases in Paris – what is the probability that three weeks later it will be (a) zero ? (b) more than 100 ?” is one such question.

As noted in Section 2.1 however, agent-based models are very demanding in terms of computational time and effort; many such studies [25, 32] use networks of 10,000 people and not larger. It is noteworthy that the two works we just cited both report on phenomena which are traditionally considered outside the ambit of lumped-parameter models; mass testing and isolation in the former study and containment of the epidemic in the latter. Our model is able to deliver almost all (an exception being for example the stable linear solution of Reference [32]) the results of the agent-based models at a minuscule fraction of the computational cost. The versatility and computational ease also mean that any newly reported information regarding say the immune response can be incorporated immediately, and the resulting case trajectories computed virtually on the same day that the study is released. This will be a huge advantage for analysing situations like spreading of variant strains or vaccination.

Finally, although we have written this Article in terms of the specific disease COVID-19, our conclusions should remain valid even after the pandemic is over. Infectious disease has always been a part of human existence, and with the advent of jetliner travel, pathogens can be carried halfway across the world in a matter of hours. With current trends continuing, it is highly probable that pandemics are here to stay. We hope that the same may be said for our model as well.

APPENDIX A. DERIVATION OF REPRODUCTION NUMBER R

As stated in the main text, the reproduction number R denotes the number of people to whom one case transmits the disease on average. There are two ways we can obtain R from (2.3), one qualitative and the second quantitative. We present both derivations below.

For the qualitative derivation, we note that the per-case spreading rate m_0 denotes the number of targets to whom an at-large case transmits the disease every day, assuming they are all susceptible. When a fraction y/N of these targets are immune, the average spreading rate will decrease by this fraction and become $m_0(1 - y/N)$. By the model, a randomly selected case is contact traced with a probability $1 - \mu_3$ in which situation s/he spreads the disease for τ_h days. With probability μ_3 the random case is untraced; there is a further probability $1 - \mu_1$ that s/he is symptomatic and spreads it for τ_2 days while there is a probability μ_1 that s/he is asymptomatic and spreads it for τ_1 days. Thus, on average a case spreads the disease for

$$T_s = (1 - \mu_3)\tau_h + (1 - \mu_1)\mu_3\tau_2 + (1 - \mu_1)\mu_3\tau_1, \quad (\text{A.1})$$

days, which is called the serial interval. Multiplying the per-day spreading rate by the serial interval gives the reproduction number

$$R = m_0 \left(1 - \frac{y}{N}\right) \left(\frac{1 + \mu_3 - 2\mu_1\mu_3}{2}\tau_2 + \mu_1\mu_3\tau_1\right). \quad (\text{A.2})$$

Here we have used the fact that $\tau_h = \tau/2$.

This derivation of R was primarily qualitative, so we now adopt a rigorous mathematical approach. Here, we attempt to obtain the stability of a constant solution $y = y^*$ to (2.3). For this, we first linearize about this point so that the logistic factor $1 - y^*/N$ can be treated as a constant. This leaves us with the structure

$$dy/dt = C [y(t) - (1 - \mu_3)y(t - \tau_h) - (1 - \mu_1)\mu_3y(t - \tau_2) - \mu_1\mu_3y(t - \tau_1)], \quad (\text{A.3})$$

where C denotes $m_0(1 - y^*/N)$. When all the delays are zero, (A.3) becomes an ODE which has the solution $y = \text{const}$. We can see that $y = \text{const}$ is also a solution of the full DDE (A.3). So, if we try a solution to the DDE of the form $e^{\lambda t}$ where λ is an eigenvalue, then one of these eigenvalues will be zero. The stability of the constant solution will be governed by the other eigenvalues – if the real part of the largest eigenvalue is negative then a constant solution will be stable while if that real part is positive then a constant solution will be unstable. The former case corresponds to termination of the epidemic in time while the latter case corresponds to growth of

the epidemic in time. When the two delays are very small, we find that the solutions of (A.1) are stable, while they become unstable if the delays are increased. We now seek the condition for the stability transition.

The characteristic equation of (A.3) is

$$\lambda = m_0 \left[1 - (1 - \mu_3) e^{-\lambda\tau_h} - (1 - \mu_1) \mu_3 e^{-\lambda\tau_2} - \mu_1 \mu_3 e^{-\lambda\tau_1} \right]. \quad (\text{A.4})$$

We can see that $\lambda = 0$ is always a solution. When the delays are identically zero, all the other eigenvalues have a real part of $-\infty$; when the delays are increased, they start approaching the imaginary axis. The transition occurs when they actually breach this axis. There are two ways this can occur – a real root crosses the axis in a saddle node bifurcation or a pair of complex roots cross the axis in a Hopf. We first try the latter case. In this case, right at the transition point, there will be one pair of roots having the value $\pm j\omega$ for some non-zero frequency ω . Substituting this into (A.4) and equating the real and imaginary parts we get

$$m_0 \left[1 - (1 - \mu_3) \cos \omega\tau_h - (1 - \mu_1) \mu_3 \cos \omega\tau_2 - \mu_1 \mu_3 \cos \omega\tau_1 \right] = 0, \quad (\text{A.5a})$$

$$m_0 \left[(1 - \mu_3) \sin \omega\tau_h + (1 - \mu_1) \mu_3 \sin \omega\tau_2 + \mu_1 \mu_3 \sin \omega\tau_1 \right] = \omega. \quad (\text{A.5b})$$

Now, (A.5a) has the structure $1 - aC_h - bC_2 - cC_1$ where C_1 denotes $\cos \omega\tau_1$ etc and a, b and c stand for the three constants involved. This implies that $aC_h + bC_2 + cC_1 = 1$. But, a, b, c are positive and satisfy $a + b + c = 1$. If they are weighted by any kind of cosines, their sum will no longer be 1, unless each of the cosines is itself 1. Hence, this must be true for (A.5a). Immediately, each of the sines in (A.5b) becomes zero, and that equation reduces to $\omega = 0$, which contradicts the starting premise of the case. Hence, (2.1) cannot undergo a Hopf bifurcation.

With Hopf ruled out, the transition to instability in (A.3) has to be through zero crossing of a single real root. At the instant of the transition, that will be a real root having the value zero. Since $\lambda = 0$ is already a root, at the transition it will become a repeated root. This means that (A.3) will have a solution proportional to t in addition to a constant solution. The stability criterion comes out by forcing that $y = t$ be a solution of (A.3); we immediately get

$$m_0 \left[1 - \frac{y^*}{N} \right] \left[(1 - \mu_3) \tau_h + (1 - \mu_1) \mu_3 \tau_2 + \mu_1 \mu_3 \tau_1 \right] = 1 \quad (\text{A.6})$$

Recalling that τ_h is defined to be half of τ_2 , the criterion reduces to

$$m_0 \left[1 - \frac{y^*}{N} \right] \left(\frac{1 + \mu_3 - 2\mu_1\mu_3}{2} \tau_2 + \mu_1 \mu_3 \tau_1 \right) = 1 \quad (\text{A.7})$$

Now, since y^* can be anything, we can just replace it with y to obtain the stability criterion at any level of the disease. We can see that the solution is stable if the LHS of (A.7) is less than 1 and unstable if it is greater than 1. This LHS is the same as what we obtained in (A.2) hence we can identify it as the reproduction number R .

During the growth phase of the disease, the doubling time can be calculated using the reproduction number R and the serial interval T_s , as follows. By definition, if there are α new cases cropping up today, then there will be $R\alpha$ new cases cropping up after T_s days. Fitting this to an exponential growth $e^{\gamma t}$ we find $\gamma = \log R/T_s$. The doubling time is $T_d = \log 2/\gamma = (\log 2 T_s)/\log R$.

APPENDIX B. CONTACT TRACING

Here we present the derivation of the equation in Section 3.2. To develop this aspect, we first assume that there is no contact tracing whatever; then (2.3) looks like

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y}{N} \right] [y(t) - (1 - \mu_1)y(t - \tau_2) - \mu_1 y(t - \tau_1)]. \quad (\text{B.1})$$

Now we have to throw away the assumption of zero NTI. For this, we introduce a new variable denoting the exposed people. In the ODE models it is denoted by E ; since we are using y for cases, we see no harm in using x for exposures instead. We define that a person transitions from x to y at the instant s/he turns transmissible, which is also the moment when s/he will first return positive if tested for the virus.

What the RHS of (B.1) denotes is actually the rate of exposures – interaction between cases and targets leads to exposures rather than cases among the latter. So, we can replace the left hand side (LHS) of (3.5) by dx/dt as a first step in going past the zero NTI assumption. Doing this we have

$$\frac{dx}{dt} = m_0 \left[1 - \frac{x}{N} \right] [y(t) - (1 - \mu_1)y(t - \tau_2) - \mu_1 y(t - \tau_1)]. \quad (\text{B.2})$$

Here, in the susceptibility probability, we must count anyone who has not been exposed yet, so it features x ; the spreading on the other hand is done by at large cases and not exposures so the counting term features y . Now let τ_4 denote the period during which a person remains exposed before turning into a case – as in all lumped parameter models, this must be an average over all cases. Then, all the people who got exposed between time t and $t + \Delta t$ will turn into cases between $t + \tau_4$ and $t + \tau_4 + \Delta t$. This implies that $\dot{x}|_t = \dot{y}|_{t+\tau_4}$. Since both x and y start from zero at $t = 0$, we can integrate this to write $x(t) = y(t - \tau_4)$ or,

$$\frac{dy}{dt} \Big|_{t+\tau_4} = m_0 \left[1 - \frac{y(t - \tau_4)}{N} \right] [y(t) - (1 - \mu_1)y(t - \tau_2) - \mu_1 y(t - \tau_1)], \quad (\text{B.3})$$

or,

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y(t - 2\tau_4)}{N} \right] [y(t - \tau_4) - (1 - \mu_1)y(t - \tau_2 - \tau_4) - \mu_1 y(t - \tau_1 - \tau_4)]. \quad (\text{B.4})$$

The difference between (B.1) and (B.4) is just a shift of the infection curves by τ_4 , which is a small quantity in relation to the epidemic's overall progress (a typical value of τ_4 is 4 days). Hence, in the baseline model we could afford to make the zero NTI assumption without significant error.

For a sophisticated contact tracing model however, a little reasoning shows that this assumption is unacceptably simplistic. For if a patient's contacts were to be rounded up as soon as they have been exposed, then they would all go into quarantine during their NTI itself and not spread the disease to anyone. In practice however, contact tracing is time-consuming – the starting patient's test results have to come in (which itself may take a day or more), authorities have to talk to the patient, find out the names and places s/he gives, call the people involved, access CCTV footage and/or card histories in public places (if permitted) etc. The authorities might be understaffed/overworked and might take time to get to the patients; the patients might not remember their entire movements accurately and might give out the story piecemeal. So, by the time the authorities get to the contacts, some will have already crossed the NTI into the transmissible asymptomatic or pre-symptomatic infection stage.

We assume here that the contact tracing takes place as follows. When a symptomatic patient reports corona to the authorities, they start identifying the people with whom s/he came into contact over the past τ_2 days. They take a time τ_3 to unearth and quarantine a fraction P_3 of the cases among these contacts. The quarantine of healthy people is an unpleasant collateral of a contract trace and does not affect our model. With this strategy

of *one-level, forward* contact tracing, let us try to formulate the equivalent of (B.2). The heuristic contact-traced fraction μ_3 of (2.3) must now get replaced by an expression cast in terms of the ground realities τ_3 , τ_4 and P_3 and the other existing parameters.

When a symptomatic case, say Mr X, just reports for quarantine, he will have spent the preceding τ_2 days unknowingly spreading virus to targets. In any lumped parameter formulation, we must assume that Mr X spreads the disease uniformly and continuously during his pre-symptomaticity period – this assumption makes sense when not one but a hundred simultaneous cases are taken into consideration. Among Mr X's targets are the person Ms Y he infected right when he turned transmissible and the person Ms Z he infected just before reporting for quarantine. When Mr X locks himself up, the former has spent τ_2 days at large while the latter has spent 0 days at large. The authorities take a further τ_3 days to obtain these contacts, at which point Ms Y has got an exposure time of $\tau_2 + \tau_3$ while Ms Z has got an exposure time of τ_3 . There are three possible scenarios now.

- Case 1 : $\tau_4 > \tau_2 + \tau_3$. In this case, Ms Y is still in NTI and hence every other contact of Mr X is in NTI as well. Those that are identified do not transmit – the average time T spent transmissible and at large by the contact traced cases is identically zero.
- Case 2 : $\tau_2 + \tau_3 > \tau_4 > \tau_3$. In this case, Ms Y has passed NTI period into transmission but Ms Z is still in NTI. Let $t = 0$ be the time when Mr X turned transmissible so that $t = \tau_2$ is when he quarantines, and let n be the total number of contacts whom he has infected with $r = n/\tau_2$ being the infection rate. By hypothesis, the contact trace occurs at $t = \tau_2 + \tau_3$; the last of the exposures who turns transmissible before being traced is the one who was exposed at $t = \tau_2 + \tau_3 - \tau_4$. The target who was exposed to Mr X at any time t prior to $\tau_2 + \tau_3 - \tau_4$ turns transmissible at time $t + \tau_4$ and thus remains transmissible and at large for a duration $(\tau_2 + \tau_3) - (t + \tau_4)$. The total time which all these cases have spent at large is

$$U = \int_0^{\tau_2 + \tau_3 - \tau_4} (\tau_2 + \tau_3 - \tau_4 - t) r \, dt = \frac{r}{2} (\tau_2 + \tau_3 - \tau_4)^2. \quad (\text{B.5})$$

Recalling that $r\tau_2 = n$ and dividing by n gives the average time spent at large by contact traced cases as

$$T = \frac{(\tau_2 + \tau_3 - \tau_4)^2}{2\tau_2}. \quad (\text{B.6})$$

If τ_3 and τ_4 are both zero (which marginally satisfies the conditions of this case), *i.e.* we make the assumptions of zero NTI and instantaneous tracing, then this reduces to $\tau_2/2$, an important consistency check.

- Case 3 : $\tau_4 < \tau_3$. In this case Ms Z turns transmissible as well before isolation. We can use a procedure similar to Case 2 above to calculate an average time for which the contact traced case remains at large. Since the situation is implausible and the algebra is tedious, we skip this exercise.

So we have found the average duration T that a contact traced case remains at large – what we now need is the absolute probability that a random case is contact traced (recall that P_3 is the *conditional* probability that a case is contact traced, *given that s/he* was exposed by a symptomatic case). This absolute probability is composed of two probabilities : (a) that the random case has been infected by a symptomatic case and not an asymptomatic one (secondary cases of asymptomatic cases do not get picked up by definition), and (b) P_3 . To calculate the probability (a), we use a self-consistency procedure.

Let this unknown probability be X . We assume, very plausibly, that the secondary contacts of every case (symptomatic or otherwise) are asymptomatic with the fraction μ_1 (as defined for the disease as a whole) and symptomatic with the fraction $1 - \mu_1$. Consider a random asymptomatic case. With probability X she has been infected by a symptomatic case; with further probability P_3 she gets picked up in the contact tracing drive. If she does, *i.e.* with probability P_3X , she spends time T at large. Otherwise, she remains at large for the entire

asymptomatic infection period τ_1 . Combining the two, on average an asymptomatic case remains at large for a time

$$T_1 = (1 - P_3X) \tau_1 + P_3XT. \quad (\text{B.7})$$

Similarly, a symptomatic case gets picked up in the drive and spreads for time T with probability P_3X ; he spends the entire pre-symptomatic period τ_2 at large otherwise. Thus, the average time spent at large by symptomatic cases is

$$T_2 = (1 - P_3X) \tau_2 + P_3XT. \quad (\text{B.8})$$

So we have a fraction μ_1 of asymptomatic cases remaining at large for time T_1 and a fraction $1 - \mu_1$ of symptomatic cases remaining at large for time T_2 . Since by the model assumptions all cases transmit equally and uniformly, the total numbers of secondary cases spawned by these two types of cases must be in the ratio $\mu_1 T_1$ to $(1 - \mu_1) T_2$. This ratio must be *the same as the ratio of the probabilities* that a random case has been infected by an asymptomatic and a symptomatic case respectively; thus we have

$$\frac{1 - X}{X} = \frac{\mu_1 [(1 - P_3X) \tau_1 + P_3TX]}{(1 - \mu_1) [(1 - P_3X) \tau_2 + P_3TX]}, \quad (\text{B.9})$$

which rearranges into a quadratic equation for X

$$(P_3T - \mu_1 P_3 \tau_1 - (1 - \mu_1) P_3 \tau_2) X^2 + (\mu_1 \tau_1 + (1 - \mu_1) [P_3 \tau_2 + \tau_2 - P_3 T]) X - \tau_2 (1 - \mu_1) = 0. \quad (\text{B.10})$$

One of its two roots, $X = X^*$, should be a number between 0 and 1 and amount to the desired probability.

Finally, we are at a stage where we can write the model equation. A random case is contact traced with probability P_3X^* , so this is the equivalent of $1 - \mu_3$ in the baseline model (2.3). The contact traced cases remain at large for duration T as calculated in the case-wise analysis. Thus, the exposure equation is given by

$$\frac{dx}{dt} = m_0 \left[1 - \frac{x}{N} \right] [y(t) - P_3X^*y(t - T) - (1 - \mu_1)(1 - P_3X^*)y(t - \tau_2) - \mu_1(1 - P_3X^*)y(t - \tau_1)], \quad (\text{B.11})$$

and the case equation follows as

$$\begin{aligned} \frac{dy}{dt} = m_0 & \left[1 - \frac{y(t - 2\tau_4)}{N} \right] \\ & \times [y(t - \tau_4) - P_3X^*y(t - T - \tau_4) - (1 - \mu_1)(1 - P_3X^*)y(t - \tau_2 - \tau_4) - \mu_1(1 - P_3X^*)y(t - \tau_1 - \tau_4)], \end{aligned} \quad (\text{B.12})$$

which is (3.5) of the main Article. The structure is very similar to (2.3) but the terms T and X^* now arise from calculations of considerable sophistication, based directly on the details of the contact tracing process. Similarly, it should also be possible (though hasslesome) to implement two-level contact tracing where contacts of contacts are also identified and grounded, and reverse tracing where the source of a freshly reporting case is also attempted to be tracked down.

APPENDIX C. TESTING WITH LIMITED SENSITIVITY

Here we show how to incorporate the effect of limited test sensitivity into the equation (11). A study [20] indicates that the ‘‘gold standard’’ RT-PCR test has a sensitivity of approximately 75 percent or lower; the antigen tests are typically less sensitive. With imperfect sensitivity, if a person is a case, the test has a probability

P of actually reporting positive and a probability $1 - P$ of coming out false negative. We assume that the tests are fully specific *i.e.* there are no false positive cases (this assumption is consistent with reality). To calculate the durations for which cases remain at large, we use the same argument as above. For the asymptomatic Mr X, this time there is a probability P that he gets caught after 4 days and a probability $1 - P$ that he remains at large for the full 7 days. Similarly for the symptomatic Ms Y, there is probability P of capture after 18/7 days and probability $1 - P$ of remaining at large for 3 days. The asymptomatic at large time is therefore

$$T_1 = 4P + 7(1 - P), \quad (\text{C.1})$$

while the symptomatic at large time is

$$T_2 = \frac{18}{7}P + 3(1 - P). \quad (\text{C.2})$$

Using these in the basic structure (2.1) we have

$$\frac{dy}{dt} = m_0 \left[1 - \frac{y}{N} \right] [y(t) - (1 - \mu_3)y(t - T) - (1 - \mu_1)\mu_3y(t - T_2) - \mu_1\mu_3y(t - T_1)], \text{ where} \quad (\text{C.3a})$$

$$T = \frac{\mu_1 T_1 + (1 - \mu_1) T_2}{2} \quad (\text{C.3b})$$

This completes the derivation.

REFERENCES

- [1] <https://cmmid.github.io/topics/covid19/>.
- [2] <https://www.imperial.ac.uk/mrc-global-infectious-disease-analysis/covid-19/covid-19-publications/6>.
- [3] <https://covid-19.bsvgateway.org/>.
- [4] D. Adak, A. Majumder and N. Bairagi, Mathematical perspective of COVID-19 pandemic: disease extinction criteria in deterministic and stochastic models, MedRxiv Article (2020) available at <https://www.medrxiv.org/content/10.1101/2020.10.12.20211201v1>.
- [5] D. Adjodah *et al.*, Decrease in hospitalizations for COVID-19 after mask mandates in 1083 US counties. MedRxiv (2020). Available at <https://www.medrxiv.org/content/10.1101/2020.10.21.20208728v1>.
- [6] M. Agrawal, M. Kanitkar and M. Vidyasagar, Modeling the spread of SARS-CoV-2 pandemic: impact of lockdowns and interventions. To appear *Indian J. Med. Res.* (2020).
- [7] J.M. Cashore *et al.*, COVID-19 mathematical modeling for Cornell's fall semester, (2020) available at [https://cpb-us-w2.wpmucdn.com/sites.coecis.cornell.edu/dist/3/341/files/2020/10/COVID_19_Modeling_Ju\\$N_1\\$5-VD6.pdf](https://cpb-us-w2.wpmucdn.com/sites.coecis.cornell.edu/dist/3/341/files/2020/10/COVID_19_Modeling_Ju$N_1$5-VD6.pdf).
- [8] M.L. Childs *et al.*, The Impact of long term non-pharmaceutical interventions on COVID-19 epidemic dynamics and control, MedRxiv Article (2020) available at <https://www.medrxiv.org/content/10.1101/2020.05.03.20089078v1>.
- [9] V. Chin *et al.*, A Case study in model failure – COVID-19 daily deaths and ICU bed utilization predictions in New York State, Arxiv Article [2006.15997](https://arxiv.org/abs/2006.15997) (2020).
- [10] A. Das, A. Dhar, S. Goyal and A. Kundu, COVID-19 : analysis of an extended SEIR model and comparison of different interventions strategies, Arxiv Article [2005.11511](https://arxiv.org/abs/2005.11511) (2020).
- [11] L. Dell'Anna, Solvable delay model for epidemic spreading : the case of COVID-19 in Italy, Arxiv Article [2003.13571](https://arxiv.org/abs/2003.13571) (2020).
- [12] A. Dhar, A Critique of the COVID-19 analysis for India by Singh and Adhikari, *ibid.* 2004.05373 (2020).
- [13] O. Diekmann,, J.A.P. Heesterbeek and J.A.J. Metz, On the Definition and computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28** (1990) 365–382.
- [14] A.W.D. Edridge *et al.*, Human coronavirus reinfection dynamics : lessons for SARS-CoV-2, *ibid.* (2020) available at <https://www.medrxiv.org/content/10.1101/2020.05.11.20086439v2>.
- [15] G. Giordano *et al.*, Modeling the COVID-19 epidemic and implementations of population-wide interventions in Italy. *Nat. Med.* **26** (2020) 855–860.
- [16] Y. Gu, Estimating true infections : a simple heuristic to measure implied infection fatality rate, (2020) available at <https://covid19-projections.com/estimating-true-infections/>.
- [17] W.O. Kermack and A.G. McKendrick, A Contribution to the mathematical theory of epidemics. *Proc. Roy. Soc. A* **115** (1927) 700–721.

- [18] R.J. Kosinski, The Influence of time-limited immunity on a COVID-19 epidemic: a simulation study. MedRxiv Article (2020) available at <https://www.medrxiv.org/content/10.1101/2020.06.28.20142141v1>.
- [19] M. Kreck and E. Scholz, Studying the course of COVID-19 by a recursive delay approach, *ibid.* (2021) available at <https://www.medrxiv.org/content/10.1101/2021.01.18.21250012v2>.
- [20] L.M. Kucirka, S.A. Lauer, O. Laeyendecker, D. Boon and J. Lessler, Variation in false-negative rate of RT-PCR tests for SARS-CoV2 by time since exposure, *ibid.* (2020) available at <https://www.medrxiv.org/content/10.1101/2020.04.07.20051474v1>.
- [21] J. Lavine, O. Bjornstadt and R. Antia, Immunological characteristics will govern the changing severity of COVID-19 during the likely transition to endemicity, *ibid.* (2020) available at <https://www.medrxiv.org/content/10.1101/2020.09.03.20187856v1>.
- [22] G. Menon, Problems with the Indian supermodel for COVID-19, available at <https://www.thehindu.com/sci-tech/science/problems-with-the-indian-supermodel-for-covid-19/article32937184.ece>.
- [23] J. Mossong, N. Hens and M. Jit, Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLOS Medicine* **5** (2008) e74.
- [24] F. Sandmann *et al.*, The Potential health and economic value of SARS-CoV-2 vaccination alongside physical distancing in the UK : transmission model-based future scenario analysis and economic evaluation, *ibid.* (2020) available at <https://www.medrxiv.org/content/10.1101/2020.09.24.20200857v1>.
- [25] S.R. Serrao *et al.*, Requirements for the containment of COVID-19 disease outbreaks through periodic testing, isolation and quarantine, MedRxiv Article (2020) available at <https://www.medrxiv.org/content/10.1101/2020.10.21.20217331v1>.
- [26] M.M. Sharma and B. Shayak, Public health implications of a delay differential equation model for COVID-19, Proceedings of KIML Workshop, KDD2020 (2020) available at https://aiisc.ai/KiML2020/papers/KiML2020_paper_77.pdf.
- [27] B. Shayak, Differential Equations – Linear Theory and Applications, available electronically at www.shayak.in/Shayakpapers/DELTA/DELTA8.pdf.
- [28] B. Shayak, M.M. Sharma and M. Gaur, A New delay differential equation for COVID-19, Proceedings of KIML Workshop, KDD2020 available at https://aiisc.ai/KiML2020/papers/KiML2020_paper_69.pdf.
- [29] B. Shayak, M.M. Sharma, R.H. Rand, A. Singh and A. Misra, A Delay differential equation model for the spread of COVID-19 submitted, available at <http://www.shayak.in/Shayakpapers/Diseases/Corona35.pdf>.
- [30] B. Shayak and R.H. Rand, Self-burnout – a new path to the end of COVID-19, MedRxiv Article (2020) available at <https://www.medrxiv.org/content/10.1101/2020.04.17.20069443v2>.
- [31] R. Singh and R. Adhikari, Age-structured impact of social distancing on the COVID-19 epidemic in India, Arxiv Article 2003.12055 (2020).
- [32] S. Thurner, P. Klimer and R. Hanel, A Network-based explanation of why most COVID-19 infection curves are linear. *PNAS* **117** (2020) 22684–22689.
- [33] C.P. Vyasarayani and A. Chatterjee, New approximations and policy implications from a delayed dynamic model of a fast pandemic. *Physica D* **414** (2020) 132701.
- [34] A. Wajnberg *et al.*, SARS-CoV-2 infection induces robust neutralizing antibody responses that are stable for at least three months, *ibid.* (2020) available at <https://www.medrxiv.org/content/10.1101/2020.07.14.20151126v1>.
- [35] L.S. Young, S. Ruschel, S. Yachuk and T. Pereira, Consequences of delays and imperfect implementation of isolation in epidemic control. *Sci. Rep.* **9** (2019) 3505.