A TWO PARAMETER CLOSED FORM SOLUTION TO KERMACK AND MCKENDRICK'S INTEGRO-DIFFERENTIAL EQUATIONS

Ted Duclos, Tom Reichert

Author affiliations:

T. Duclos, Ted Duclos Advisors, St. Augustine, Florida, ORCID: 0000-0002-5623-2623, email address: tedduclos56@gmail.com

T. Reichert, Entropy Research Institute, Portland, Oregon. ORCID: 0000-0003-3006-5385

Author contributions: T. Duclos and T. Reichert: Conceptualization, Writing – Original draft preparation, Writing – Review & Editing. T. Duclos: Methodology, Validation. T. Reichert: Data Investigation.

Funding: This work was funded internally by Ted Duclos Advisors and ERI.

Declarations of interest: none.

Word counts: Abstract, 194 words; text, 8803 words

Key Words

Kermack and McKendrick integro-differential equations

Closed form solution

Epidemic management

Identification of outbreaks

Network Concepts

Abbreviations used in this text

KMES	Kermack McKendrick integro-differential Equation Solution
K&M	Kermack and McKendrick
RCO	Rate of Change Operator
SIR	Susceptible–Infectious–Recovered
CI	Confidence Interval

Abstract

A closed form solution of the full Kermack and McKendrick (K&M) integro-differential equations (Kermack and McKendrick 1927), called the KMES, is presented and verified. Intuitively, the KMES states that the total case count in an epidemic is the integral of the product of the probability rate of transmission and the number of infectious contacts from the epidemic start to the current point in time.

This solution was derived by pursuing the logical consequences of a construct that combines network concepts with the recognition that the infectiousness of an individual depends on their exposure to susceptible persons. The KMES has two parameters: the probability rate of disease transmission, and the basic reproduction number. The proof of the solution uses the weighted averages of K&M's rate parameters, and a restatement of the proper integral limits in their recovery function.

Practical analytical expressions, including novel expressions for the effective reproduction number, time to peak in new infections, and the final size flow directly from the solution. Notably, the latter two expressions project phenomena antithetical to currently accepted epidemic concepts. Using early COVID -19 pandemic data from six different countries to estimate the model parameters, the KMES accurately projects case data from the COVID-19 pandemic in the following 60-day period with R^2 values > 0.78. In addition, the projection of a typical individual's contagiousness, derived from the KMES, closely approximates the time course of viral shedding as measured in infected persons.

Introduction

Modern epidemiological modeling is rooted in the Kermack and McKendrick (K&M) epidemic model (Kermack and McKendrick 1927). Since its publication, well over 10,000 authors have referenced this paper as a foundational starting point. However, as Diekmann points out in his insightful essay (2022 pg. 8), "…an incessant community-enforced misconception is that the paper is just about the very special case, the S-I-R (Susceptible, Infected, Recovered) model." In the essay, he also discusses his suspicion that, despite the numerous references, the citing authors rarely read the 1927 paper. His concluding remarks decry the situation and request that researchers carefully study the 1927 paper, a "…true gem…in which tremendous wisdom lies hidden." (Diekmann 2022 pg. 9).

A solution to K&M's full integro-differential equations would unlock further wisdom, but the widespread use of SIR approximations (e.g., Breda et al 2021, Diekmann 2022 and Brauer 2008) and the publication of a discrete time model of the integro-differential equations (Diekmann 2021), imply that such a solution remains elusive. Absent a solution, researchers have made prodigious efforts to extract wisdom from the SIR approximations and produced the well-known formulas for the final size of epidemics, the notion of "flatten the curve" (DiLauro et al 2021), and various other refinements (e.g., Brauer 2008, Breda et al 2021, Hethcote 2000).

Unfortunately, approximations can incorporate unrealistic assumptions. For instance, SIR approximations inherently assume that the affected population is well-mixed; a supposition Brauer (2008, pg. 27) labelled as "unrealistically simple". Lacking analytical alternatives,

advances in the use of approximations tend to focus on minimizing the effect of the associated assumptions.

One alternative approach, network-based stochastic models (e.g., Newman 2002, Diekmann et al 1998, and Youssef 2011) assumes that epidemic dynamics can be projected by summing the collective probabilities that individuals will be infected by a limited number of contacts; thereby avoiding the well-mixed assumption. These widely investigated approaches have produced estimations of outbreak threshold conditions and the final size of epidemics. However, while they are laudatory steps towards realistic characterizations of interactions, they do not yield the analytical expressions which make closed-form solutions so appealing and useful.

Fortunately, in his essay, Diekmann (2022) also offers a clue to the form a solution should take by stating that the infection incidences are "…the probability per unit of time of (a) susceptible (person) becoming infected…times the susceptible subpopulation". Since this also requires that there must be contacts between susceptible and infectious persons for incidences to occur, we can anticipate that an expression for the time course of the total incidences should be interpretable as the integral, over time, of the product of the probability rate of infection transmission and the number of infectious contacts with susceptible people.

Inspired by Diekmann's comments, we studied K&M's 1927 paper; and, by applying network model concepts, found a pathway to a closed form solution which we call the KMES.

Our approach, laid out below, frees the analysis from the assumptions inherent in the SIR approximations and reveals fundamental attributes of K&M's equations:

1) In Section 1 we present a closed form solution to the K&M integro-differential equations and demonstrate its veracity, both graphically and analytically. The solution derivation leans heavily on the principle that the infectiousness of a contagious person depends on contact with susceptible people; and the proof of the solution relies upon both the weighted average of K&M's recovery rate parameter and the proper integration limits in their probability of infectiousness function. Analytical expressions characterizing epidemic dynamics derive from the solution.

2) In Section 2, using the solution and data compiled from six countries during the COVID-19 pandemic, we demonstrate that we can deduce a probability rate of transmission and population interaction parameter from early pandemic data. With these two parameters, the solution not only accurately projects the course of the COVID-19 pandemic in the six sampled countries, but also provides an estimate of the contagiousness of individuals.

3) In Supplements 2 and 3, we derive and demonstrate the use of additional expressions which can aid public health officials to diagnose, control, and even end an epidemic.

Section 1: A Solution to the Kermack and McKendrick Equations

We first present an overview and elucidation of K&M's integro-differential equations, envision the epidemic as a network and then use this to find a solution to these equations.

Kermack and McKendrick's Integro-Differential Equations

The terms used in K&M's model are: N_p, the total population; A_p the area that contains N_p and $S(t)$, the subpopulation that has not yet become infected. The ever-infected population is subdivided as: $N(t,\theta)$, the persons who are currently or have previously been infected (the cases); $I(t,\theta)$, the individuals currently infectious within $N(t,\theta)$; and $\mathrm{R}(t,\theta)$ the recovery of individuals within $N(t,\theta)$; where θ is the time since infection of these subpopulations. We call $S(t)$, $N(t,\theta)$, $I(t,\theta)$, and $\mathrm{R}(t,\theta)$ the "population variables", and remark that a population variable integrated over θ from $0 \to t$ is a function of time alone (e.g., $I(t) = \int_0^t I(t,\theta)\, d\theta$). As well, K&M defined $\varphi(t,\theta)$ as the probability rate of new infection transmission per susceptible density and infectiousness; and $\psi(t,\theta)$ as the probability rate of recovery per infectiousness.

Given these definitions, $N(t) = I(t) + R(t)$, and $\mathrm{S}(0) = N_p$. We further assume that N_P is a constant, $I(0) = N(0) = 1$, $R(0) = 0$, and that, once infected, people cannot become reinfected. Also, since θ has the units of time, $\Delta\theta = \Delta t$.

K&M's integro-differential equations, written in terms of these definitions, are:

$$\frac{dS(t)}{dt} = -\frac{dN(t)}{dt} = -\frac{\mathrm{S(t)}}{A_p}\left(\int_0^t A(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + A(t,t)I(0)\right), \qquad (1)$$

$$I(t) = \int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0), \qquad (2)$$

$$\frac{dR(t)}{dt} = \int_0^t C(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + C(t,t)I(0), \qquad (3)$$

$$N_p = S(t) + N(t), \qquad (4)$$

$$N(t) = I(t) + R(t), \tag{5}$$

where $B(t,\theta) = e^{-\int_0^\theta \psi(t,a)da}$, $C(t,\theta) = \psi(t,\theta)B(t,\theta)$, and $A(t,\theta) = \varphi(t,\theta)B(t,\theta)$. Although K&M (1927) wrote their equations in terms of subpopulation densities, we have purposefully chosen, without altering their result, to depict Equations 1-3 in terms of subpopulation totals.

K&M derived these equations by imagining $N(t)$ as the sum at time t of incremental subpopulations, $\frac{dN(t-\theta,0)}{dt}$, which we shall here call "θ-groups", infected at time $t-\theta$. They subsequently tracked the remaining infectiousness of each θ-group through both time and θ using the function, $B(t,\theta) = e^{-\int_0^\theta \psi(t,a)da}$, which is the probability, based on the time since infection, that a person in $N(t)$ remains infectious. Equations 1-3 were then constructed by evaluating the residual infectiousness, $B(t,\theta)\frac{dN(t-\theta,0)}{dt}$, within each θ-group, summing each θ-group's contribution to the changes in the susceptible and recovered subpopulations, and to the total infectious subpopulation.

Specifically, in Equation 1, K&M's term describing the contribution to the change in the susceptible population (i.e., the negative of the change in new infections), $-\frac{\mathrm{S(t)}}{A_p}\varphi(t,\theta)B(t,\theta)\frac{dN(t-\theta,0)}{dt}$, is each θ-group's residual infectiousness, scaled by the susceptible density, and the probability rate of transmission. Similarly, in Equation 3, the term describing the contribution to the change in the recovered populations, $\psi(t,\theta)B(t,\theta)\frac{dN(t-\theta,0)}{dt}$, is a fraction of the residual infectiousness as determined by the probability rate of recovery. By integrating these θ-group contributions over θ, K&M created equations for the rate of change of

both $S(t)$ and $R(t)$, and completed the model by summing the residual infectiousness of the θ-groups to find $I(t)$ in Equation 2.

Consistent with Diekmann's statement, a solution to Equation 1 will be the integral, over time, of the probability of transmission times the number of infectious contacts with susceptible people. Problematically, however, K&M included terms for infectious and susceptible people in Equation 1 because both need to be non-zero for new infections to occur. This approach implies a well-mixed population, which, if assumed to be true, makes finding a solution to the system of equations difficult. Nevertheless, as revealed in the following analysis, this difficulty can be overcome by adopting analytically tractable groupings of infectious contacts and parameters.

The Epidemic as a Network

The path to a solution begins by envisioning the population affected by the epidemic as a network represented by a connected graph with N_P persons/vertices, and with edges that represent the contacts between the N_P persons with an infection transmission probability of one. We assume that all contacts other than the edges have an infection transmission probability of zero, hence, the mean number of the edges per person is the self-weighted mean degree of this network minus 1, and is, by definition, the basic reproduction number, R_0.

If an epidemic progresses within N_P, this conceptualization of R_0 translates as each member of $N(t)$ being associated with a group of contacts of average size R_0; groups which we label "N-groups". Persons within $N(t)$ are not members of their own N-group and each N-group

consists of the people, who, if not yet infected, their associated contagious person will ultimately infect.

This notion of R_0 is consonant with a conventional concept of R_0, in which $R_0 = pc(t_1 - t_2)$ (Diekman et al 2013, pg. 4); where p is the probability a contact will become infected, c is the expected contact rate, and $t_1 - t_2$ is the time period when p is non-zero. In our approach, since R_0 is the number of contacts that will become infected, $p = 1$ during the period $t_1 - t_2$; and, consequently, $c(t_1 - t_2)$ is the number of people who will become infected. We do not need to specify either the period $t_1 - t_2$ or the contact rate c hidden within R_0 to derive our solution; however, later in our analysis, we apply a probability rate of transmission to capture the dynamics of infection transmission.

Of course, not everyone who is infected can transmit the disease. Therefore, we can distinguish three subgroups of infected people. These are:

1) Latently Infectious: People who have been infected but are not yet shedding virus.

2) Actively Infectious: People who are shedding the virus and are in contact with susceptible people.

3) People who may be shedding the virus but have no contact with susceptible people are part of the recovered group, $R(t)$.

Persons in groups 1 and 2 constitute $I(t)$ and are collectively known as infectious. Persons in group 2 are infectious and contagious, and persons in group 3 are contagious, but not infectious. All three groups, together with the previously infected, but now recovered group, comprise the ever-infected population, $N(t)$.

We assume in our construction, that infectors remain in durable potentially infectious contact with the people they have infected. Consequently, the infectiousness of infectors, but not the degree within the network, diminishes with each new infection they cause; and, in symmetry, the extent to which they are unable to infect others, their recovery, increases. We also assume in our analysis that R_0 stays constant, no two N-groups are identical, and, for simplicity, that the disease does not have a latent period.

The Solution

Employing this N-group perspective, we find the solution by first totaling the changes within the N-groups in a manner analogous to the way that K&M summed the changes in the θ-groups. These changes equate to the sums of the changes in K&M's θ-groups; and the solutions to the resulting N-group based differential equations also solve K&M's equations.

Since the size of each N-group on average equals R_0 and remains constant, the totality of the contacts in the N-groups can be divided into two time-varying portions: $R_{eff}(t)$, the average number of contacts within each N-group that are susceptible and remain to be infected; and $R_N(t)$, the average number of contacts that are previously infected people. Given that the

susceptible persons within the N-groups are, by definition, in contact with an infectious person, $\frac{N(t)R_{eff}(t)}{N(t)R_0}$ is the probability that a member of $N(t)$ is infectious (i.e., $\frac{R_{eff}(t)}{R_0} = \frac{I(t)}{N(t)}$). Thus,

$$I(t) = N(t)\ \frac{R_{eff}(t)}{R_0}. \tag{6}$$

We highlight here that the product of $\frac{R_{eff}(t)}{R_0}$ and $N(t)$ yields the total residual infectiousness of the N-groups. Because this directly corresponds to K&M's calculation of $I(t)$ in Equation 2 by summing the residual infectiousness of the θ-groups, we assert that Equation 6 is the form of Equation 2 for the N-groups.

If we define $\frac{K_T}{R_0}$ as the probability rate of transmission between susceptible and contagious contacts, assume it to be constant across the population, and assume K_T is a quantity associated with the disease, then,

$$\frac{dS(t)}{dt} = \frac{dR_{eff}(t)}{dt}N(t) = -K_T N(t)\frac{R_{eff}(t)}{R_0}, \tag{7}$$

an expression that we assert is the N-group form of Equation 1. Notably, since the term $R_{eff}(t)$ equates to the number of infectious contacts taking place, in contrast to K&M's Equation 1, we do not need to use the product of the infectious and susceptible subpopulation magnitudes in this expression to indicate the level at which these contacts are occurring.

Since $R_{eff}(0) = R_0$, integrating and exponentiating Equation 7 produces,

$$R_{eff}(t) = R_0 e^{-\frac{K_T}{R_0}t}, \tag{8}$$

Applying Equation 6 to Equation 8, and defining $F_i(t)$ as the fraction of $N(t)$ that is infectious,

$$\frac{I(t)}{N(t)} = F_i(t) = e^{-\frac{K_T}{R_0}t}. \tag{9}$$

Continuing the derivation, we remark that, $R_N(t) = R_0 - R_{eff}(t)$, and by Equation 8,

$$R_N(t) = -R_0(e^{-\frac{K_T}{R_0}t} - 1). \tag{10}$$

Changes in $R_{eff}(t)$ negatively covary with $R_N(t)$, and $\frac{dR_{eff}(t)}{dt}N(t) = -\frac{dN(t)}{dt}$, thus,

$$\frac{dR_N(t)}{dt} = \frac{dN(t)}{N(t)dt}, \tag{11}$$

which is simply a restatement of Equation 7. Since $R_N(t)$ goes from zero to $R_N(t)$, as $N(t)$ goes from 1 to $N(t)$, integrating both sides of Equation 11 produces,

$$R_N(t) = \ln\,(N(t)). \tag{12}$$

Exponentiating Equation 12 and substituting Equation 10 for $R_N(t)$, we find,

$$N(t) = e^{-R_0(e^{-\frac{K_T}{R_0}t}-1)}. \tag{13}$$

This is a closed form expression for $N(t)$. Combining Equation 13 and 9 we obtain,

$$I(t) = e^{-R_0(e^{-\frac{K_T}{R_0}t}-1)-\frac{K_T}{R_0}t}, \tag{14}$$

which is a closed form expression for $I(t)$. Subtracting Equation 14 from Equation 13 generates a closed form expression for $R(t)$:

$$R(t) = (1 - e^{-\frac{K_T}{R_0}t})\, e^{-R_0(e^{-\frac{K_T}{R_0}t} - 1)}. \tag{15}$$

We call Equations 13 through 15 the KMES (**K**ermack and **M**cKendrick **E**quation **S**olution) and maintain that they are a solution to Equations 1 through 5. We present graphical and analytical proofs supporting this claim in the next two subsections.

The logic of the KMES can be seen by first observing that the term $e^{-\frac{K_T}{R_0}t}$ is the probability that any member of $N(t)$ is infectious. Therefore, the expressions for $I(t)$ and $R(t)$ are the product of $N(t)$ times $e^{-\frac{K_T}{R_0}t}$ and $1 - e^{-\frac{K_T}{R_0}t}$ respectively. Additionally, as anticipated by Diekman's statement, the expression for $N(t)$ itself is the integral, up to the current time t, of the product of the probability rate of transmission $\frac{K_T}{R_0}$ and the number of infectious contacts $I(t)R_0$.

Graphical proof

To show, graphically, that the KMES projects the same time course of $N(t), I(t)$, and $R(t)$ as K&M's integro-differential equations (Equations 1-5), we must first restate Equations 1-5 in terms of K_T and R_0 without using the KMES. With this end in mind, we note that since the left sides of Equations 1 and 7 as well as the left sides of Equations 2 and 6 are equivalent, we can employ Equations 6 through 11, and, in essence, apply our network assumptions and N-group summations to K&M's equations to make the restatement.

To begin, if we apply Equation 6 and the derivative of Equation 4 to Equation 7, then,

$$\frac{dN(t)}{dt} = K_T I(t), \tag{16}$$

If we next define the time-based weighted averages of $\varphi(t,\theta)$ and $\psi(t,\theta)$ as:

$$\hat{\varphi}(\mathrm{t}) = -\frac{A_p \frac{dS(t)}{dt}}{S(t)I(t)} = \frac{\int_0^t A(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + A(t,t)I(0))}{\int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0))}, \text{ and} \tag{17}$$

$$\hat{\psi}(t) = \frac{\frac{dR(t)}{dt}}{I(t)} = \frac{\int_0^t C(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + C(t,t)I(0)}{\int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0)}, \tag{18}$$

we then find that,

$$K_T = \hat{\varphi}(t)\frac{\mathrm{S(t)}}{A_P}. \tag{19}$$

Equation 19 illustrates that the change in perspective from θ-groups to N-groups, typified by Equation 7, subsumes $S(t)$ into K_T. It also shows that a constant probability rate of transmission between infectors and their eventual recipients means that $\hat{\varphi}(t)$ and $S(t)$ are inversely proportional, a result congruent with their units.

Then, because $R_0 = R_{eff} + R_N$, by Equation 6 and Equation 5,

$$R(t) = N(t)\frac{R_N(t)}{R_0}. \tag{20}$$

Differentiating both sides of Equation 20 produces,

$$\frac{dR(t)}{dt} = \frac{dN(t)}{dt}\frac{R_N(t)}{R_0} + \frac{N(t)}{R_0}\frac{dR_N(t)}{dt}, \tag{21}$$

which, using Equations 10, 11, and 16, becomes,

$$\frac{dR(t)}{dt} = (K_T - K_T e^{-\frac{K_T}{R_0}t} + \frac{K_T}{R_0})I(t), \tag{22}$$

and, by Equation 18 and 22, we find that,

$$\hat{\psi}(t) = K_T - K_T e^{-\frac{K_T}{R_0}\,t} + \frac{K_T}{R_0}. \tag{23}$$

To give an intuitive sense to Equation 22, we note that $\frac{1}{R_0}$ is the portion of the new infections, $K_T I(t)$, that diminishes the infectiousness of the infectors as they incrementally infect their N-groups; and the term, $K_T - K_T e^{-\frac{K_T}{R_0}\,t}$, is the portion of the existing infections, $I(t)$, that have recently recovered from the disease.

By Equations 17 and 18, we can assume that $\hat{\varphi}(t) = \varphi(t, \theta)$ and $\hat{\psi}(t) = \psi(t, \theta)$, without affecting the value of either Equation 1 or 3; therefore, Equations 22 and 3 are equivalent. Also, since $\frac{dN(t)}{dt} = -\frac{dS(t)}{dt}$, the substitution of Equations 19, 17, and 2 into Equation 16 illustrates the equivalency between Equation 16 and Equation 1. Subsequently, if we subtract Equation 22 from Equation 16, we produce an expression equivalent to the derivative of Equation 2:

$$\frac{dI(t)}{dt} = (K_T e^{-\frac{K_T}{R_0}t} - \frac{K_T}{R_0})I(t). \tag{24}$$

In summary, by applying our network assumptions, embodied in the N-group based equations, and employing the weighted averages, the K&M integro-differential equations can be restated as the following system of equations:

$$\frac{dN(t)}{dt} = K_T I(t), \tag{16}$$

$$\frac{dI(t)}{dt} = (K_T e^{-\frac{K_T}{R_0}t} - \frac{K_T}{R_0})I(t), \text{ and} \tag{24}$$

$$\frac{dR(t)}{dt} = (K_T - K_T e^{-\frac{K_T}{R_0}t} + \frac{K_T}{R_0})I(t). \tag{22}$$

$$N_p = S(t) + N(t), \tag{4}$$

$$N(t) = I(t) + R(t), \tag{5}$$

Since this system of equations is equivalent to K&M's integro-differential equations, and was found <u>without</u> using the KMES, we can use it to graphically prove that the KMES is the solution to K&M's equations. To this end, we used the parameters listed in Table 1 in a fourth order Runga-Kutta algorithm to simulate Equations 16, 24, and 22, and then plotted these alongside the KMES in Figure 1. As illustrated in the figure, the KMES replicates the simulation with an exceedingly small error; thereby proving that the KMES solves K&M's equations.

Analytical proof

As confirming proof, we can insert the KMES into Equations 1-5 and show analytically that it solves them. We begin by solving Equation 2, a task which uses the outcome, derived in Supplement 1, that the infectiousness probability function can be written as:

$$B(t, \theta) = e^{-\int_{t-\theta}^{t} \widehat{\psi}(t) da}. \tag{25}$$

As is also explained in Supplement 1, to preserve the value of the integrals in Equations 1 – 3, when using $\widehat{\psi}(t)$, the integral limits in $B(t, \theta)$ must be the interval from $t - \theta$ to t.

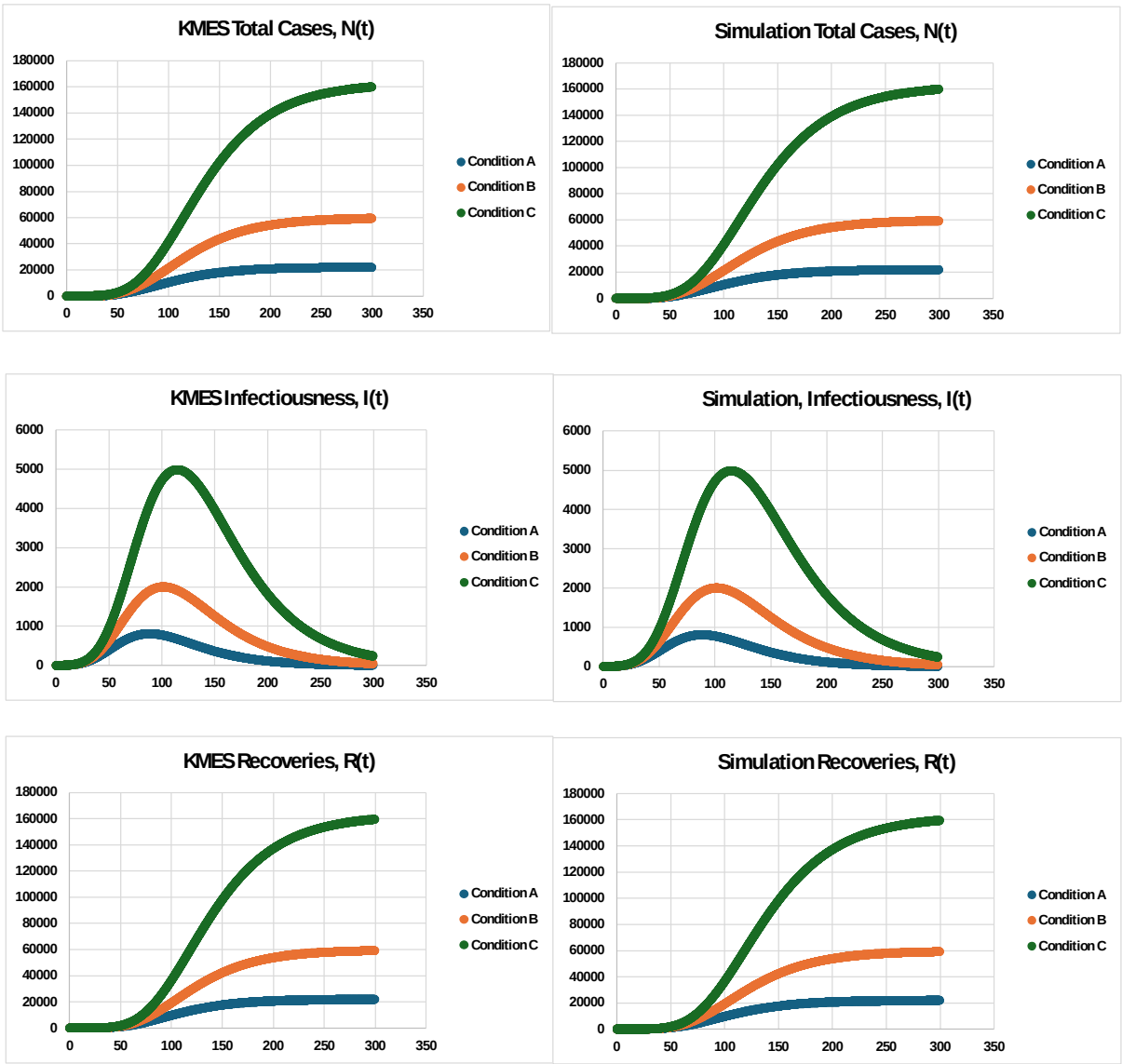

Figure 1. Calculations of $N(t)$, $I(t)$, and $R(t)$ using the KMES (Left Side; Equations 13-15), and Simulation of Equations 16, 24, and 22 (Right Side). The parameter values found in Table 1 were used to generate the plots. The projections appear in separate plots because the Mean Absolute Percent Error between them is less than 5E-7, and if they were plotted together, then the top plot would completely obscure the underlying plot.

Table 1. Values of the parameters used to create Figure 1

Condition A	Time step=	0.1	KT=	0.26	Ro=	10	I(0)=	1	Np=	40000000
Condition B	Time step=	0.1	KT=	0.26	Ro=	11	I(0)=	1	Np=	40000000
Condition C	Time step=	0.1	KT=	0.26	Ro=	12	I(0)=	1	Np=	40000000

By substituting the right side of Equation 23 into Equation 25, we obtain,

$$B(t,\theta) = e^{-K_T\theta - R_0\left(e^{-\frac{K_T}{R_0}t} - e^{-\frac{K_T}{R_0}(t-\theta)}\right) - \frac{K_T}{R_0}\theta}, \text{ and} \tag{26}$$

$$B(t,t) = e^{-K_T t - R_0\left(e^{-\frac{K_T}{R_0}t} - 1\right) - \frac{K_T}{R_0}t}. \tag{27}$$

Then, by differentiating both sides of Equation 13 and using $t-\theta$ for time, we find,

$$\frac{dN(t-\theta,0)}{dt} = I(0)K_T e^{-R_0(e^{-\frac{K_T}{R_0}(t-\theta)} - 1) - \frac{K_T}{R_0}(t-\theta)}. \tag{28}$$

Substituting these three expressions into Equation 2, we have,

$$I(t) = \int_0^t e^{-K_T\theta - R_0\left(e^{-\frac{K_T}{R_0}t} - e^{-\frac{K_T}{R_0}(t-\theta)}\right) - \frac{K_T}{R_0}\theta} K_T e^{-R_0\left(e^{-\frac{K_T}{R_0}(t-\theta)} - 1\right) - \frac{K_T}{R_0}(t-\theta)} d\theta +$$

$$e^{-K_T t - R_0\left(e^{-\frac{K_T}{R_0}t} - 1\right) - \frac{K_T}{R_0}t}. \tag{29}$$

Equation 29 is a formidable expression, nevertheless, we <u>can</u> evaluate the integral and simplify the resultant expression to prove that Equation 14 is the solution for $I(t)$ in Equation 2.

$$I(t) = e^{-R_0(e^{-\frac{K_T}{R_0}t} - 1) - \frac{K_T}{R_0}t}.$$

Equations 1 and 3 can be rewritten using the equalities $\hat{\varphi}(t) = \varphi(t,\theta)$, $\hat{\psi}(t) = \psi(t,\theta)$, and $\frac{dN(t)}{dt} = -\frac{dS(t)}{dt}$ as,

$$\frac{dN(t)}{dt} = \frac{\mathrm{S(t)}}{A_p}\hat{\varphi}(t)\left(\int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0)\right), \text{ and} \tag{30}$$

$$\frac{dR(t)}{dt} = \hat{\psi}(t)\left(\int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0)\right). \tag{31}$$

By Equation 2, the highlighted expressions on the right-hand side of both Equation 30 and 31 equal $I(t)$. Substituting the proven expression for $I(t)$ along with Equations 19 and 23 into Equations 30 and 31, then integrating, we find the rest of the solution,

$$N(t) = e^{-R_0\left(e^{-\frac{K_T}{R_0}t}-1\right)}, \text{ and}$$

$$R(t) = \left(1 - e^{-\frac{K_T}{R_0}t}\right) e^{-R_0\left(e^{-\frac{K_T}{R_0}t}-1\right)}.$$

Thus, the KMES is a closed form solution to K&M's integro-differential equations.

Expressions Describing an Epidemic

Because it is a closed form solution, expressions useful in epidemic description and management emerge from the KMES. We illustrate a few of these here.

Epidemic Direction

Using the expression for $R_{eff}(t)$, Equation 8, to rewrite Equation 25 produces,

$$\frac{dI(t)}{dt} = (R_{eff}(t) - 1)\frac{K_T}{R_0}\mathrm{I(t)}, \tag{32}$$

Therefore, if $R_{eff}(t) > 1$, the epidemic is growing, and reciprocally, if $R_{eff}(t) < 1$, the epidemic is declining. Hence, if K_T and R_0 are evaluable, then Equation 8 can assist officials in making public health decisions. In support of this application, in Section 2 we describe the methods to extract K_T and R_0 from epidemic data.

Time to Peak New Infections and Peak Size

Setting $R_{Eff}(t) = 1$ in Equation 8 yields the time when both $I(t)$ and $\frac{dN(t)}{dt}$ peak,

$$t_{peak} = \frac{R_0 \ln(R_0)}{K_T}, \tag{33}$$

Significantly, Equation 33 projects that when social interventions are stronger (smaller R_0), the time to the peak will <u>always</u> be shorter; the converse of the SIR models' projections of this relationship, known as "flatten the curve". Furthermore, Equation 33, when substituted into Equation 14, generates an equation for projecting the peak value of $I(t)$,

$$I(t)_{Peak} = \frac{I(0)e^{(R_0-1)}}{R_0}. \tag{34}$$

Final Size

In the limit as $t \to \infty$, Equation 13 simplifies to an expression for the final size,

$$N(\infty) = e^{R_0}, \tag{35}$$

Therefore, if R_0 remains large enough for long enough, the KMES projects that $N(t)$ <u>can</u> equal N_P. The prerequisite for this to occur is,

$$R_0 > \ln\,(N_P). \tag{36}$$

This result differs from K&M's (1927) conclusion that the entire susceptible population cannot be infected; a result that has achieved a position of prominence within the epidemiological modelling community. Nevertheless, since Equation 35 derives from the KMES, and the KMES solves Equations 1 through 5, we must conclude that the K&M integro-differential equations do, in fact, allow for an entire population to become infected.

K&M (1927) did not solve their equations, and therefore could not derive Equation 35. Rather, their conclusion that an entire population cannot be infected rested upon their assumption that the expression, $\int_0^\infty \varphi(t,\theta)B(t,\theta)d\theta$, remains finite for all time. This assumption, however, is mathematically unsupportable, because in Equation 1, if $\frac{dN(t)}{dt} > 0$ (i.e. new infections are occurring) and $t \neq \infty$, then as the number of susceptible people goes to zero, both $B(t,\theta)\frac{dN(t-\theta,0)}{dt}$ and $B(t,t)I(0)$ remain finite and greater than zero, and hence, $\varphi(t,\theta)$ trends to infinity. Therefore, $\int_0^\infty \varphi(t,\theta)B(t,\theta)d\theta$ can become unbounded and their conclusion was unfounded. In retrospect, had K&M adopted our concept of infectiousness and grouped $S(t)$ and $\varphi(t,\theta)$ within a K_T – like parameter, they could have avoided this misstep.

Relationship between the KMES and SIR final size expressions

From Brauer (2008), the final size equation derived from the SIR approximation is,

$$\ln\,(\frac{S(0)}{S(\infty)}) = R_0(1 - \frac{S(\infty)}{K}), \tag{37}$$

where $K \approx S(0)$. This is a well-known expression from epidemic theory, and we offer here an explanation of how, in view of Equation 35, to choose the value of $S(0)$ so that Equation 37 yields the proper final size.

By the definitions of $S(t)$, and $N(t)$,

$$S(0) = N(\infty) + S(\infty). \tag{38}$$

If we substitute this into Equation 37, we obtain:

$$\ln\left(\frac{N(\infty)+S(\infty)}{S(\infty)}\right) = R_0\left(1 - \frac{S(\infty)}{S(0)+1}\right). \tag{39}$$

We then set $N(\infty) = e^{R_0}$ in Equation 40, assume $1 + \frac{S(\infty)}{S(0)+1} \approx 1$, and solve for $S(\infty)$ to obtain,

$$S(\infty) = \frac{e^{R_0}}{e^{R_0}-1} \approx 1. \tag{40}$$

Therefore, if $S(0)$ is much larger than one and equals $e^{R_0} + 1$, then Equations 35 and 37 yield the same result.

We can generalize this finding by first acknowledging that Equation 37 is only valid when $S(\infty) > 0$ because the SIR equation it is derived from, $\frac{dS(t)}{dt} = -\frac{\beta I(t)S(t)}{N_p}$, can only be solved for $S(\infty)$ if $S(\infty)$ is assumed to be greater than zero. Therefore, since the value of $S(0)$ is an input to Equation 37, when modeling an epidemic in a population of size N_A, if $N_A > e^{R_0}$, then $S(0)$ can always be chosen such that it equals $e^{R_0} + 1$; and, with this choice, Equation 37 yields the same final size as Equation 35. For those situations where $N_A \leq e^{R_0}$, then equation 37 is not valid and Equation 35 projects that the entire population will be infected.

This analysis does not claim that Equations 35 and 37 are equivalent. Rather, it demonstrates how, with a judicious choice of $S(0)$ in Equation 37, we can correct Equation 37 to obtain the same final size as Equation 35. Without this correction, in cases where the final size is more than one person less than N_p, Equation 37 projects a final size that is too large.

Section 2: Projections of Pandemic Data Using the KMES

If we differentiate both sides of Equation 13, divide by $N(t)$, rearrange using F_i, and take the natural log of both sides, we find an expression useful for making projections,

$$RCO = ln\left(\frac{\frac{dN(t)}{dt}}{N(t)}\right) = \ln(F_i(t_1)K_T) - \frac{K_\mathrm{T}}{R_0}(t - t_1), \tag{41}$$

where $t_1 < t$. We have labelled this quantity the "Rate of Change Operator" (RCO) because it is a measure of the rate of change of $N(t)$ per person within $N(t)$. With the RCO, we can extract the values of $F_i(t_1)K_T$ and $\frac{K_\mathrm{T}}{R_0}$ from case data; and, since Equation 13 may be transformed to,

$$N(t) = N(t_1)e^{\frac{-K_T F_i(t_1)}{\frac{K_\mathrm{T}}{R_0}}\left(e^{-\frac{K_T}{R_0}(t-t_1)}-1\right)}, \tag{42}$$

we can use Equations 41 and 42 to project the total and new cases during an epidemic.

To demonstrate this process, we applied Equation 42 to data recorded in six different countries during the initial stages of the COVID-19 pandemic (Roser, et al. 2021), and created the plots in Figure 2. These plots show that the RCO trends were linear both before and shortly

after the date of the imposition of containment actions (arrows in the plots), indicating that $\frac{K_T}{R_0}$ remained constant in those periods. By fitting Equation 41 to short, nine data point portions early in the straight segments of each country's RCO time series, we then estimated values of both $\frac{K_T}{R_0}$ and $F_i(t_1)K_T$ for these countries and tabulated them in Table 2.

Table 2. Social containment parameters used to model total cases and new daily cases of infection for different countries (Roser et al 2021).

	$\frac{K_T}{R_0}$	$ln(F_i(t_1)K_T)$	$N(t_1)$	Date range for RCO fit
South Korea	0.24	−1.58	3,736	March 1–March 9
USA	0.076	−1.39	46,136	March 23–March 31
Sweden	0.036	−2.47	5,320	April 1–April 9
Italy	0.080	−1.93	31,506	March 17–March 25
Spain	0.09	−2.11	65,719	March 27–April 4
New Zealand	0.20	−0.89	155	March 24–April 1

Parameters from linear fit of rate of change operator (RCO) data in Figure 2. Slope = $\frac{K_T}{R_0}$; intercept = $ln(F_i(t_1)K_T)$; t_1 is the first day of time range used. All dates are in 2020.

Figure 3 depicts the KMES projections of the daily total case data using the parameters from Table 2 in Equation 42. Note, these are <u>not</u> curve fits, yet each projection matches the actual time series of the total cases with an $R^2 > 0.78$ for the 60 days after the first date (t_1) of the date range in Table 2. Likewise, the projected values of the daily <u>new</u> cases estimated from the data in Figure 3 appear in Figure 4. The R^2 of these projections range from 0.04 (Sweden) to 0.89 (Italy); and the times of the peaks closely mirror the observed peak positions.

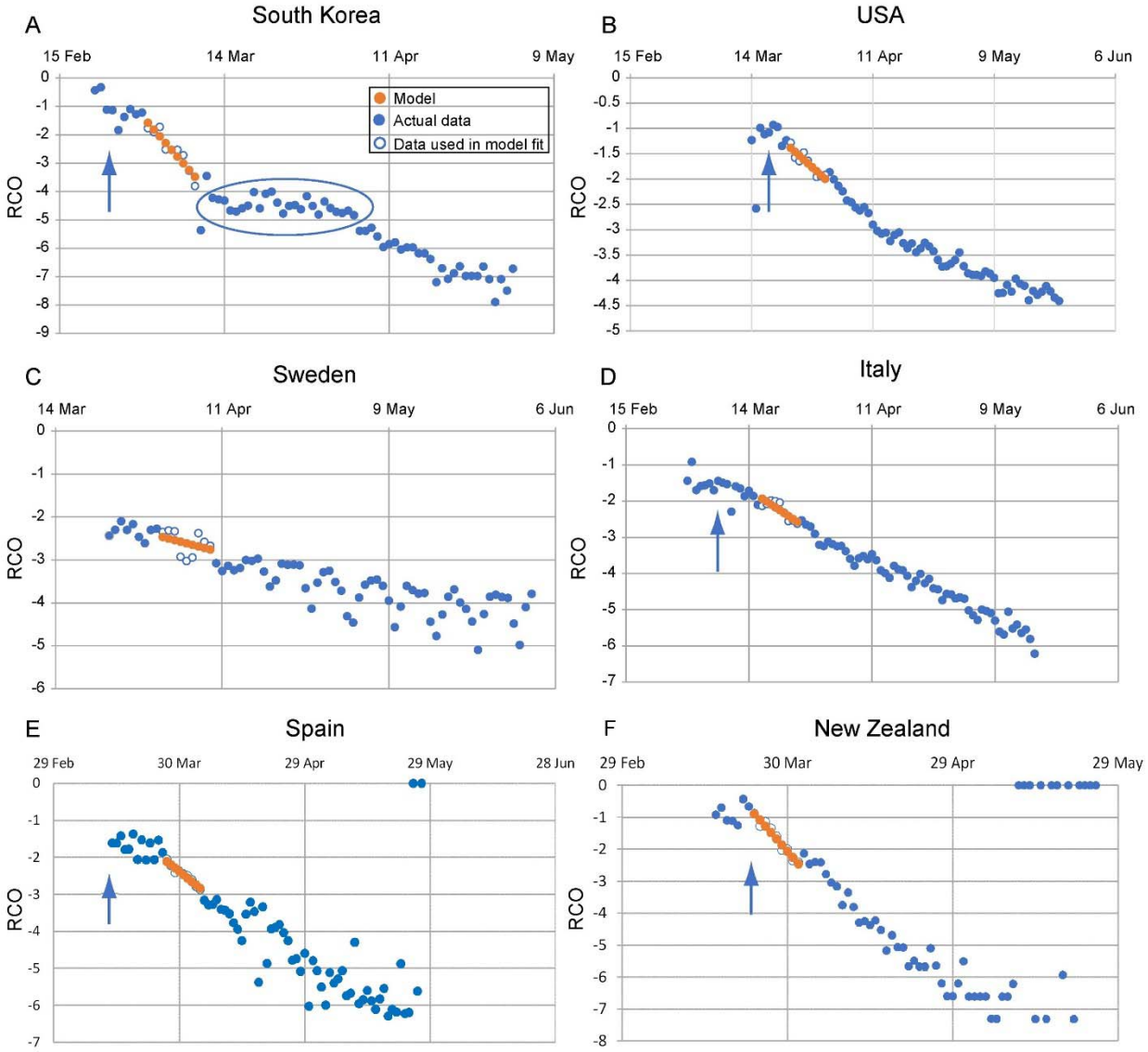

Figure 2. Rate of change operator (RCO) curves for COVID-19 cases in various countries. An epidemic can be described by a piecewise linear model using the RCO (Equation 41). A short segment of orange dots in each graph is a linear fit to the corresponding points (blue/white circles) in the observed data. The slopes of these dotted-line segments are the values of $\frac{K_T}{R_0}$ as tabulated in Table2. In some countries, RCO curves changed markedly soon after the date containment measures were implemented (arrows): **A)** South Korea, February 21; (the oval highlights a departure of the observed data from the RCO slope, indicating failures in, or relaxations of, social distancing); **B)** USA, March 16; **C)** Sweden did not implement any specific containment measures, so the short line fit was begun on April 1, the date when the slope of the RCO curve first became steady. **D)** Italy, March 8; **E)** Spain, March 14; **F)** New Zealand, March 25. All dates are in 2020.

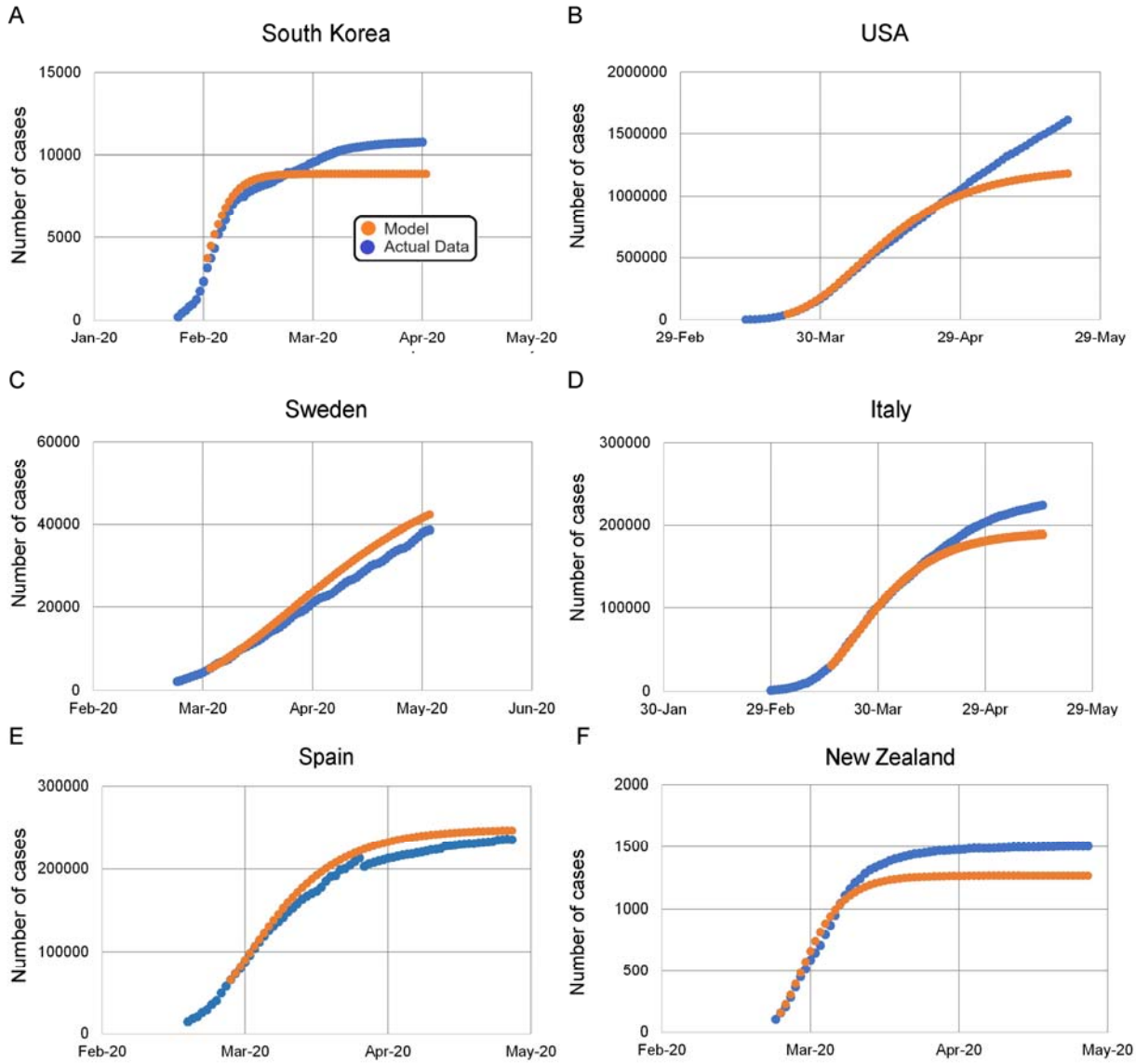

Figure 3. KMES model projections for daily total case counts. A) South Korea; **B)** USA; **C)** Sweden; **D)** Italy; **E)** Spain; and **F)** New Zealand. Dots are daily data points observed from (blue) or calculated (orange) for each country. Table 2 lists the KMES model parameters used in these plots. We found these parameter values by fitting the RCO over the date ranges listed in Table 2 and shown in Figure 2. $R^2 > 0.78$ for the model projections in all countries for the 60 days after the first date in the date range shown in Table 1. These 60 days are: South Korea, March 1-April 30; USA, March 23-May 22; Italy, March 17–May 16; Spain, March 27-May 26; New Zealand, March 24-May 23. Sweden April 1- May 31. We explain the deviation of the model from the data in the USA, panel (**B**), after April in Supplement 2. All dates are in 2020.

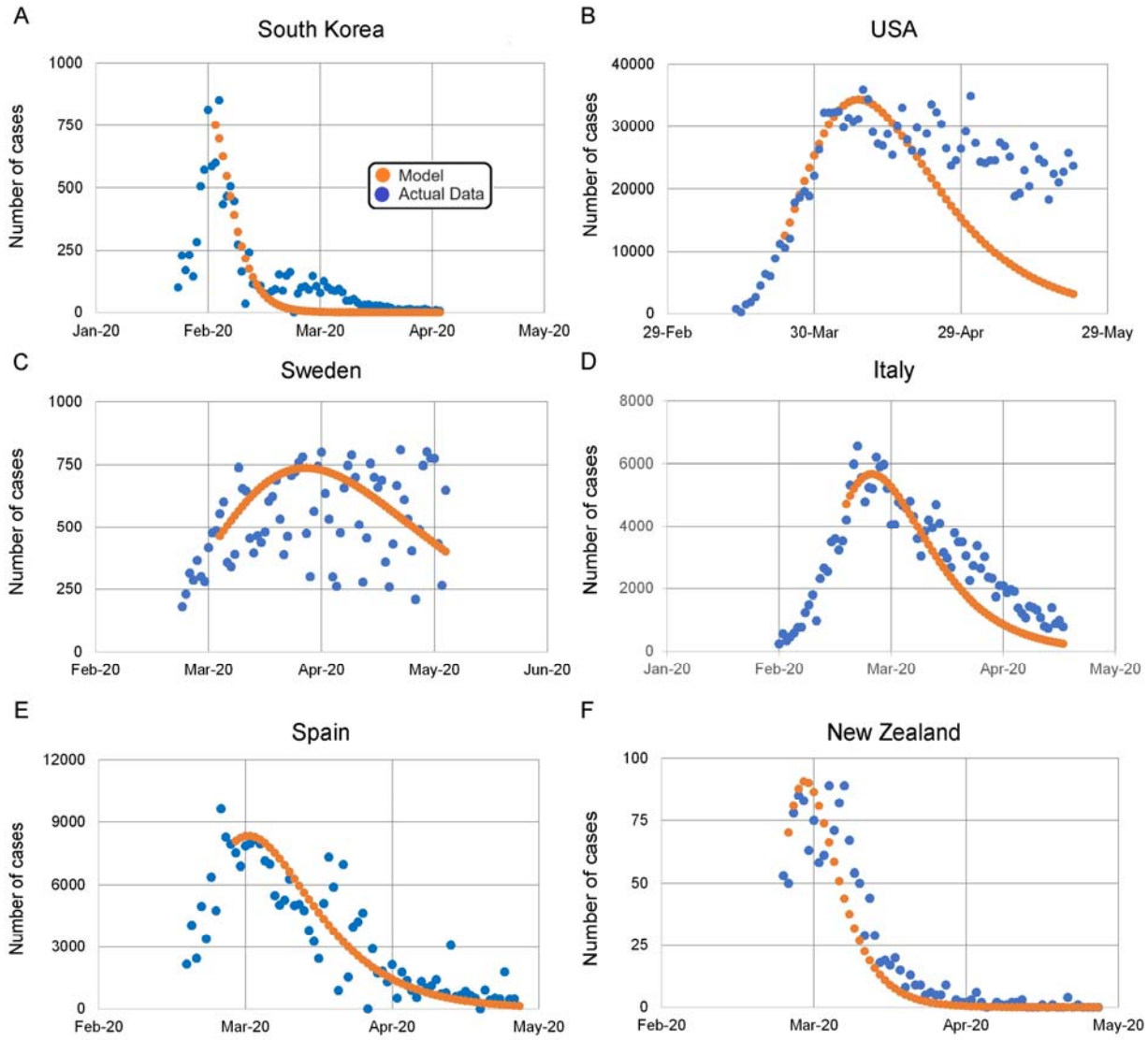

Figure 4. KMES model projections for number of new daily cases. A) South Korea, $R^2 = 0.86$; **B)** USA, $R^2 = 0.32$; **C)** Sweden, $R^2 = 0.04$; **D)** Italy, $R^2 = 0.89$; **E)** Spain, $R^2 = 0.81$; and **F)** New Zealand, $R^2 = 0.85$. The orange dotted line is the model in all panels. The blue dots are the daily observations from each country. The. R^2 values are between the model and the data, across countries for the 60 days after the date of the first date in the date range shown in Table 2.

Contagiousness

Since $B(t,t)I(0)$ is the time varying infectiousness of the initial infectors and depends upon contact with susceptible persons, the time course of contagiousness is calculable by eliminating the effect of the susceptible persons from this infectiousness. Therefore, in an imaginary scenario, if we assume that $R_0 = 1$ (i.e., the initially infected person will infect, on average, only one other person), the normalized contagiousness of the initially infected person is,

$$Normalized\ Contagiousness = \frac{B(t,t)}{R_{Eff}(t)} = e^{-(e^{-K_T t}-1)-K_T t}, \qquad (43)$$

an expression which, due to its sole dependency on K_T, supports the assumption that K_T is a disease parameter.

Estimation of K_T

To evaluate Equation 43, we must find a find a suitable value for K_T. In pursuit of this goal, we first assume that, early in the epidemic, there is a relationship between R_0 and the population density. Moreover, because the population of a country typically is only physically present within ~1% of the land in a region (Ritchie and Roser 2019), we can then define an "effective area" parameter, $A_1(t)$, where,

$$\frac{A_1(t)}{100} = Actual\ area\ traversed\ by\ a\ person.$$

If we assume that an individual's infectious mobility extends over an average effective area per unit of time, we can then write an expression in terms of this area and define $A_1(t)$ as,

$$A_1(t) = \lim_{\Delta t \to 0} \int_t^{t+\Delta t} A_{1r}(t)dt, \tag{44}$$

where $A_{1r}(t)$ equals the effective area per unit time that an individual infectiously inhabits. If we assume that changes in the value of $A_1(t)$ occur slowly in time, we can write a new expression for $\frac{K_T}{R_0}$,

$$\frac{K_T}{R_0} = \frac{K_T A_p}{A_1 N_p}, \tag{45}$$

where N_p is the entire population of the region containing the infection, A_p is the area of the region, $\frac{N_p}{A_p}$ is the population density, and A_1 is a constant.

Substituting Equation 45 into Equation 13 and solving for $-K_T t$, yields,

$$G(N(t)) = \frac{A_1 N_p}{A_p} \ln\left(1 + \frac{\ln(N(t))}{-\frac{A_1 N_p}{A_p}}\right) = -K_T t. \tag{46}$$

Since, other than K_T and A_1, the quantities in Equation 47 are available from independent sources, we can use this equation to determine the values of K_T and A_1. This procedure has two steps: first, we use an iterative process to find the value of A_1 that produces the best fit of Equation 46 to the available data; and second, we evaluate K_T as the slope of this best fit line.

Table 3 lists the available early pandemic data for our six sample countries. Using this data, through a process of iteration, we produced Figure 5 and we found that the best fit ($R^2 = 0.96$) occurs when $A_1 = 0.48$ km^2. Consequently, by the definition of A_1, the estimated <u>actual</u>

area traversed by a person is estimated to be 4800 m^2. This is a sensible result since 4800 m^2 = 80 m x 60 m, and is a plausible area for a person to traverse in a day.

In Figure 5, the line has a fitted slope of 0.26, the putative value of K_T, with a 95% confidence interval (CI) from 0.4 to 0.13. This well fit line is strong support for both assumptions that K_T is a parameter of the disease, and that K_T and A_1 were constant across the sampled countries in the initial stages of the pandemic.

Table 3. Initial COVID-19 pandemic data and population densities for various countries ((Roser et al 2021), case and date data; (Worldometers 2021, population density data)

	Date first case reported	Calculation date	Days between the first case reported and the calculation date	Cases on calculation date, N(t)	Population density (people/km^2)
South Korea	22 Jan	21 Feb	30	204	527
USA	22 Jan	19 Mar	57	13,663	36
Sweden	1 Feb	7 Mar	35	179	25
Italy	31 Jan	24 Feb	24	229	206
Spain	1 Feb	13 Mar	41	5,232	94
New Zealand	28 Feb	19 Mar	20	28	18

All dates are in 2020.

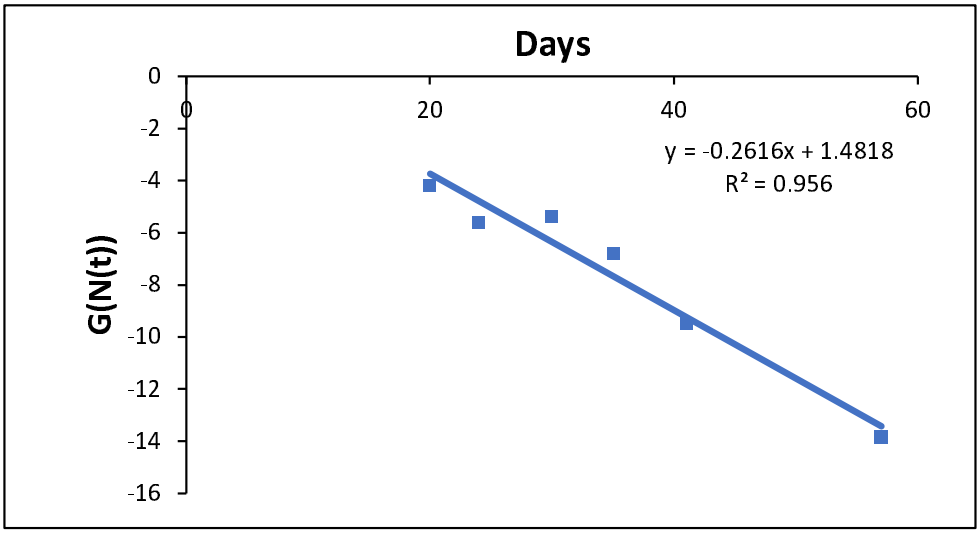

Figure 5. **Verification that K_T may be the same for all countries**. The data from Table 3 is plotted using Equation 46 and $A_1 = 0.48\ km^2$. Each data point corresponds to a different country. The fitted value of K_T is the negative of the slope of the line, and $K_T \approx 0.26$ everywhere along the line.

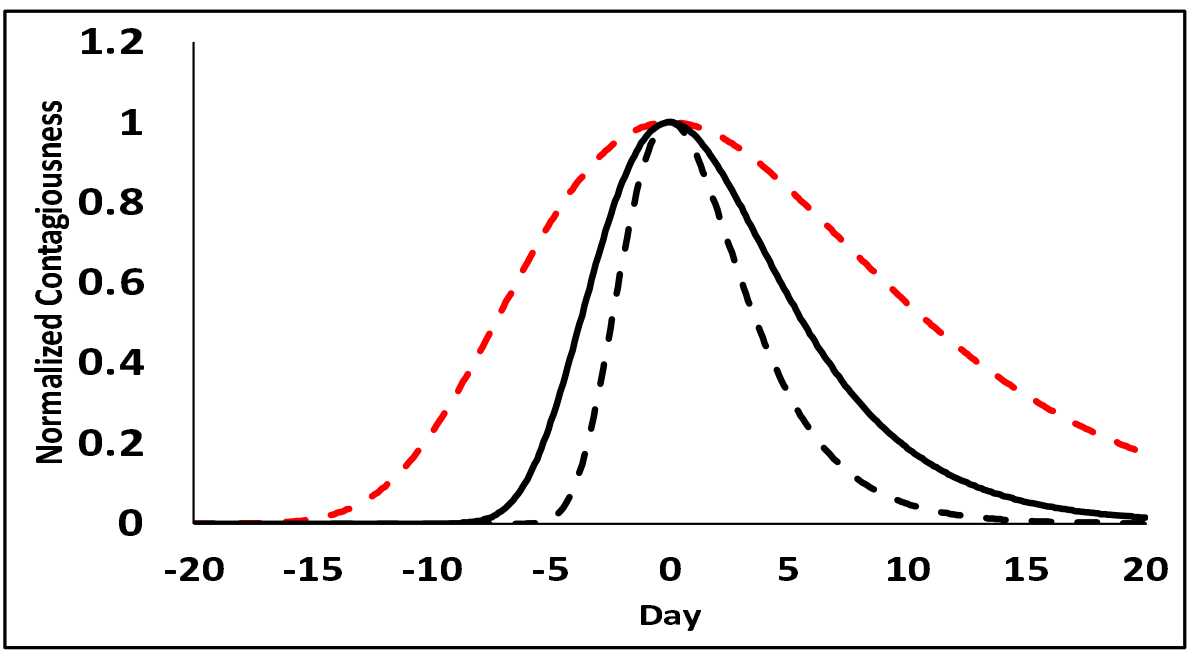

Figure 6. Representation of the normalized contagiousnesss of an initial infected person with one contact with Covid 19. The solid line was generated using the value of $K_T = 0.26$ in Equation 43; the black dashed line $K_T = 0.40$; and the red dashed line, $K_T = 0.13$; the fitted and the upper and lower 95% confidence levels, respectively, for K_T. The contagiousess begins before time zero because the initial individual was infected before the start of the epidemic.

Figure 6 is a plot of the contagiousness (Equation 43) created using the fitted and 95% CI values for K_T. The solid curve in the figure suggests that an individual infected with COVID-19 becomes contagious beginning 5 days before the peak, and declines to minimal contagiousness 10 days after the peak. This timetable corresponds with reported values of viral shedding measured in individual patients (Puhach et al 2023, Figures 2 and 3).

Discussion

In response to Diekmann's call for action, the "wisdom" we found in the Kermack and McKendrick integro-differential equations appears substantial. Foremost, we offer that the KMES obviates the need to use any approximation to the integro-differential equations. Secondly, the KMES has an intuitive probabilistic form which appears to enable the accurate projection of reported pandemic data dynamics and offers a pathway to new analytical expressions for characterizing and managing epidemics.

The analysis hinges on the N-group perspective, a convenient, and equivalent, regrouping of the parameters in K&M's integro-differential equations. Using R_0 as the point of reference, this perspective employs K&M's original concept of capturing epidemic dynamics by tracking and summing the creation and recovery of infections, but, in contrast to K&M's approach, it does so in a manner that produces resolvable expressions.

These expressions subsume the susceptible total within K_T; a favorable formulation which originates from the insight that a contagious individual is infectious in proportion to their contact with the susceptible people that they will infect. Inspired by the N-group perspective, the

use of K_T circumvents the assumption of a well-mixed population and emphasizes that the rate of new infections must be directly dependent on the contemporaneous infectiousness.

In tandem, the N-group perspective enables us to easily track the total infectiousness of the groups and tie this total to the quantity of ever infected persons and their recovery. Aiding in the proof of the solution, the replacement of K&M's two rates, $\varphi(t, \theta)$, and $\psi(t, \theta)$, by their weighted averages eliminates the variable θ and creates simpler, time-dependent differential equations. Taken altogether, these insights do not change the character of K&M's equations; rather, by adopting a more practicable method to sum the population variables, enabled by a change in perspective, they make the equations solvable.

We underscore that the closed form of the KMES rests squarely on the assumption that K_T and R_0 are constant for the duration of the epidemic. This assumption, though plausible, is an important limitation on the KMES's utility. For example, in Supplements 2 and 3 we hypothesize that any large divergence between the KMES projections and the measured data in Figure 3 may be due to fluctuations in the population interactions (and thus R_0) that drove the outcomes in the early portions of the COVID-19 pandemic. A deeper understanding of the nature of the cause-and-effect relationship between R_0 and the true population interactions could be gained by using an independent measure of population behavior in our solution as a proxy for R_0, but this can only be studied if we have a solution wherein K_T and R_0 are not required to be constant.

As a second example, a disease with a latent period could manifest itself by affecting the value of K_T, potentially requiring it to vary in time. We did not explore the analytical

consequences of this case; however, we could develop the solution further by first separating the population of latent infections from the contagious infections within $I(t)$. Then, by exploring the effect of the latent period on the solution parameters, we could determine whether the latent period can be found early in an epidemic and/or ignored once the epidemic has fully developed. Both would be useful outcomes.

A future publication will address the complexities of modelling variation in social response during epidemics with the presentation of an explicit solution to the K&M equations for the circumstance in which K_T and R_0 do vary with time. However, due to the complexity of allowing temporal variation, we have restricted this manuscript to the insight and opportunity that derives from a closed form solution.

The hybrid path combining network concepts and deterministic equations, that we used to derive the KMES, hints that herein may lie the potential to develop more sophisticated hybrids. Further exploration of the means to specify R_0 within types of networks representative of measured physical interactions between people may well lead to even more accurate forecasts of epidemic progressions. Such extensions could be useful when applied to influenza epidemic data and data from a myriad of other diseases, including RSV and the next SARS virus redux. Hybrid approaches such as these could also reduce the computational burden associated with models that employ high numbers of network nodes by allowing the calculation of results for portions of the network using analytical expressions rather than simulations.

Our analysis can be further improved by systematic address of the enormous amount of case data now available from the COVID-19 pandemic. This added analysis should refine the process required to estimate the key parameter, K_T, including variations in time, with mutations of the infectious agent and with local genetic variations in the affected population. In the inevitable future epidemics, this could substantially improve the analyses aimed at establishing the actions people and governments need to take to achieve target values of R_0.

These potential improvements notwithstanding, the closed form solution of the KMES supports and explains the findings of other researchers. For example, Equation 13 is a Gompertz equation, and as Onishi et al (2021) highlighted, the COVID-19 epidemic time course in many countries was well fit by a Gompertz model. These authors did not offer a basic principles argument for this finding. The KMES does.

Furthermore, the KMES points to a simple method to assess whether an outbreak is occurring. Embodied in R_{eff}, this offers an additional quantitative tool to guide public health officials' decision making. We illustrate this in Supplement 2.

Two conclusions, which are in contradistinction to conventional epidemic concepts, derive from the KMES. The first, an expression for the time to the peak number of new cases, Equation 33, shows that if people interact less often and decrease R_0, the peak number of infections will be much lower, and will occur much earlier, the antipode of this relationship in SIR models. In support of our conclusion, we submit that the KMES correctly and intuitively projected much lower and earlier peaks in the two countries, New Zealand and South Korea,

which imposed strong containment strategies compared to the other four sample countries (Figure 4),

The second heterodoxic conclusion, the final size expression in Equation 35, asserts that the K&M equations allow for the entire population to be infected, and lies in opposition to the conventional expression for herd immunity. We rest our arguments on the mathematical proof that the KMES solves the K&M equations, and that the final size expression derives from the KMES. We also believe that K&M's placement of $S(t)$ within Equation 1 as a stand-alone term, instead of being grouped with $\varphi(t, \theta)$ to form a single parameter, was unnecessary and indirectly lead to their (and subsequent authors') unfounded claims concerning the final size.

As a last point, and as added insight into the KMES, we highlight that the expression,

$$I(t) = e^{K_T t} B(t, t) I(0), \tag{47}$$

is an alternative form of the solution for $I(t)$. This equation states that the input of infections, $I(0)$, equates to the time varying output of infectiousness, $I(t)$, when multiplied by a step response function, $e^{K_T t} B(t, t)$. Therefore, in addition to our probabilistic interpretations, an epidemic can be cast as a dynamic system responding to a step input amplified by a function dependent on the disease and population behavior; an obvious-in-retrospect mathematical characterization.

References

Brauer, F. Compartmental Models in Epidemiology. In: Brauer, F., van den Driessche, P., Wu, J. (eds) Mathematical Epidemiology. Lecture Notes in Mathematics, 2008, vol 1945. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-78911-6_2

Breda D, Diekmann O, de Graaf W F, Pugliese A, Vermiglio R On the formulation of epidemic models (an appraisal of Kermack and McKendrick). Journal of Biological Dynamics, 2021, 6:sup2, 103-117, DOI: 10.1080/17513758.2012.716454

Campbell C. South Korea's health minister on how his country is beating coronavirus without a lockdown. TIME. https://time.com/5830594/south-korea-covid19-coronavirus/ (2020). Published April 30, 2020. Accessed February 9, 2021.

Di Lauro F, Berthouze L, Dorey MD, Miller JC, Kiss IZ. The Impact of Contact Structure and Mixing on Control Measures and Disease-Induced Herd Immunity in Epidemic Models: A Mean-Field Model Perspective, Bulletin of Mathematical Biology (2021) 83:117 https://doi.org/10.1007/s11538-021-00947-8

Diekmann O, European Communications in Mathematical and Theoretical Biology 2022 No. 24, pgs. 7-9

Diekmann O, De Jong M C M, Metz J A J, A Deterministic Epidemic Model Taking Account of Repeated Contacts between the Same Individuals Journal of Applied Probability, Vol. 35, No. 2. (June 1998), pp. 448-462

Diekmann O, Heesterbeek H, and Britton T, Mathematical Tools for Understanding Infectious Disease Dynamics, Princeton University Press (2013), ISBN 978-0-691-15539-5

Diekmann O, Othmer H G, Planque R, Bootsma, M C. The discrete-time Kermack–McKendrick model: A versatile and computationally attractive framework for modeling epidemics. PNAS 2021 Vol. 118 No. 39 e2106332118. https://doi.org/10.1073/pnas.2106332118

Elassar A, This is where each state is during its phased reopening. CNN. https://edition.cnn.com/interactive/2020/us/states-reopen-coronavirus-trnd. Published May 27, 2020. Accessed February 9, 2021.

Field A, New Zealand isn't just flattening the curve. It's squashing it. The Washington Post. https://www.washingtonpost.com/world/asia_pacific/new-zealand-isnt-just-flattening-the-curve-its-squashing-it/2020/04/07/6cab3a4a-7822-11ea-a311-adb1344719a9_story.html. Published April 7, 2020. Accessed February 9, 2021.

Hethcote HW. The Mathematics of Infectious Diseases. SIAM Review December 2000, Vol. 42 No. 4, 599-653

Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics, Proc R Soc Lond A. 1927, 115(772):700–721. https://doi.org/10.1098/rspa.1927.0118

Newman, M E J, Spread of epidemic disease on networks, PHYSICAL REVIEW E **66**, 016128, 2002

Ohnishi A, Namekawa Y, Fukui T. Universality in COVID-19 spread in view of the Gompertz function. Prog. Theor. Exp. Phys. 2020, 12:123J01. https://doi.org/10.1093/ptep/ptaa148

Puhach, O., Meyer, B. & Eckerle, I. SARS-CoV-2 Viral Load and Shedding Kinetics. *Nat Rev Microbiol* **21**, 147–161 (2023). https://doi.org/10.1038/s41579-022-00822-w

Ritchie H, Roser M. Land use. Our world in data. https://ourworldindata.org/land-use Published September 2019. Accessed February 9, 2021.

Roser M, Ritchie H, Ortiz-Ospina E, et al. Coronavirus pandemic (COVID-19). Our World in Data. https://ourworldindata.org/coronavirus. Published December 2, 2021. Accessed December 2, 2021.

Worldometers. https://.worldometers.info. Published August 21, 2021. Accessed August 21, 2021

Youssef M, Scoglio C, An individual-based approach to SIR epidemics in contact networks, Journal of Theoretical Biology, Volume 283, Issue 1, 21 August 2011, Pages 136-144, https://doi.org/10.1016/j.jtbi.2011.05.029

Supplement 1: Infectiousness Probability Function $B(t,\theta)$ using $\widehat{\psi}(t)$

To derive $B(t,\theta)$ in terms of $\hat{\psi}(t)$ we first illustrate the steps to find $B(t,\theta)$ in terms of $\psi(t,\theta)$, followed by the use of $\hat{\psi}(t)$ at a step when it <u>can</u> be substituted for $\psi(t,\theta)$, and then repeat the remaining steps to determine $B(t,\theta)$ in terms of $\hat{\psi}(t)$.

Since $B(t,\theta)$ comprises the recovery from infectiousness of each θ-group at time, t, we begin its derivation with the change in the total infectiousness,

$$\frac{\mathrm{dI(t)}}{dt} = \frac{\mathrm{dN(t)}}{dt} - \frac{\mathrm{dR(t)}}{dt}, \tag{S1-1}$$

Then, as a reminder, Equation 3 is,

$$\frac{dR(t)}{dt} = \int_0^t \psi(t,\theta)B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + \psi(t,\theta)\mathrm{B(t,t)I(0)}. \tag{S1-2}$$

which, by the definition of $I(t,\theta)$, and since $B(t,t)(I(0) + \frac{dN(0,0)}{dt}) = I(t,t)$ becomes,

$$\frac{dR(t)}{dt} = \int_0^t \psi(t,\theta)I(t,\theta)d\theta. \tag{S1-3}$$

Equation S1-1 can be discretized to,

$$\Delta I(t) = \Delta N(t) - \Delta R(t), \tag{S1-4}$$

and rewritten using the discrete form of Equation S1-3, the definition $I(t) = \int_0^t I(t,\theta)\, d\theta$, and the equality $\Delta t = \Delta\theta$, to yield,

$$I(t) - \sum_{\theta=0}^{t-\Delta t} I(t-\Delta t,\theta) = \Delta N(t) - \sum_{\theta=0}^{t-\Delta t} I(t-\Delta t,\theta)\,\psi(t-\Delta t,\theta)\Delta t, \text{ and} \tag{S1-5}$$

$$I(t) = \Delta N(t) + \sum_{\theta=0}^{t-\Delta t} I(t-\Delta t,\theta)(1-\psi(t-\Delta t,\theta)\Delta t). \tag{S1-6}$$

We initiate a useful sequence by expanding the right-hand side of Equation S1-6 with one term,

$$I(t) = \Delta N(t) + I(t-\Delta t,0)(1-\psi(t-\Delta t,0)\Delta t)$$

$$+\sum_{\theta=0}^{t-2\Delta t} I(t-2\Delta t,\theta)(1-\psi(t-2\Delta t,\theta)\Delta t)(1-\psi(t-\Delta t,\theta+\Delta t)\Delta t). \tag{S1-7}$$

Then, continually writing out terms on the righthand side, we generate the following sequence,

$$I(t) = \Delta N(t) + I(t-\Delta t,0)B(t,\Delta t) + I(t-2\Delta t,0)B(t,2\Delta t) \ldots I(0,0)B(t,t), \tag{S1-8}$$

where,

$$B(t,\theta) = \prod_{a=0}^{\theta-\Delta t}(1-\psi(a+t-\theta,a)\Delta t). \tag{S1-9}$$

Applying K&M's assumption that ψ only depends on θ and taking the limit as Δt goes to zero, we find that $B(t,\theta) = e^{-\int_0^\theta \psi(t,a)da}$, K&M's expression for $B(t,\theta)$.

The remaining derivation of $B(t,\theta)$ using $\hat{\psi}(t)$ starts with Equation 18,

$$\hat{\psi}(t) = \frac{\frac{dR(t)}{dt}}{I(t)} = \frac{\int_0^t C(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + C(t,t)I(0)}{\int_0^t B(t,\theta)\frac{dN(t-\theta,0)}{dt}d\theta + B(t,t)I(0)}, \tag{S1-10}$$

which, by the definition of $I(t,\theta)$, and since $B(t,t)(I(0)+\frac{dN(0,0)}{dt}) = I(t,t)$, means that,

$$\frac{dR(t)}{dt} = \int_0^t \hat{\psi}(t) I(t,\theta) d\theta. \qquad \text{(S1-11)}$$

Comparing Equations S1-3 and S1-11 we see that we can assume $\psi(t,\theta) = \hat{\psi}(t)$ for all θ and t without affecting either the values of the integrals or of $\frac{dR(t)}{dt}$. Next, since the right most summation in Equation S1-5 is the discrete form of the integral on the right hand of Equation S1-3, we can replace $\psi(t,\theta)$ with $\hat{\psi}(t)$ in Equation S1-5 and repeat the steps from Equation S1-5 through S1-8. The resulting product sequence for $B(t,\theta)$ is then,

$$B(t,\theta) = \prod_{a=t-\theta}^{t-\Delta t} (1 - \hat{\psi}(a)\Delta t), \qquad \text{(S1-12)}$$

which, in the limit as $\Delta t \to 0$, becomes $B(t,\theta) = e^{-\int_{t-\theta}^{t} \hat{\psi}(t) da}$, the relationship we seek.

This important result shows that, when using $\hat{\psi}(t)$, the integral limits in $B(t,\theta)$ must be the time interval from $t-\theta$ to t; the same time interval covered in Equation S1-9 as $\Delta t \to 0$ and a goes from 0 to $\theta - \Delta t$. Using this interval ensures we use the proper values of $\hat{\psi}(t)$ in $B(t,\theta)$.

Supplement 2. Controlling epidemics early

The relationships derived from the KMES in Section 1 illustrate that strong and early intervention is not only critical for control at any time, but especially early in the epidemic and at every genetic change in the disease agent. For instance, Equation 35 quantifies that the final

number of individuals infected in an epidemic depends *exponentially* on the number of people with whom each person interacts.

Real-world country data exemplifies these consequences. Both South Korea and New Zealand enacted strong and early interventions compared to other countries (Campbell, C 2020; Field, A 2020), as reflected by their higher $\frac{K_T}{R_0}$ values (Table 2). These interventions led to earlier peaks in new cases and to far fewer total cases than in other countries (Figures 2 and 3) in the first months of the pandemic. The peak number of new cases in both South Korea and New Zealand was 90–99% lower than in other countries, a validation of the statement in the KMES that strong interventions lead to *exponentially* better outcomes.

In the USA, interventions started on March 16 began to have an effect around March 23, 2020 (Figure 1B), and the number of new cases on March 23, 2020 (Roser et al 2021) was 46,136. Using the values of ln $(F_i(t_1)K_T(\mathrm{t}))$ and $\frac{K_T}{R_0}$ in Table 2, Equation 35 predicts that had these remained constants, the ultimate number of cases would have been approximately 1.22 million. On the other hand, if public officials and the population had attained the same intervention, on March 10, when there were 59 times fewer (782) cases (Roser et al 2021), then the model predicts that the number of cases would also have been 59 times lower, or 20,725. Thus, earlier action could have reduced the number of projected cases by more than 98%.

Of course, the projected estimate of 1.22 million total USA cases would only have occurred if the populace had sustained the effectiveness of the interventions launched on March

16. Unfortunately, a marked reduction in effective interventions occurred in parts of the USA in mid-April, well before the official reopening of the economy (Elassar 2020). This caused a second surge in new cases in late April and is why the observed data and the KMES predictions diverge in Figures 2B and 3B.

As shown in Section 1, Equation 33 furnishes an estimate of the time to the peak of new cases, t_{peak}. By using the data from Table 2 in Equation 33, we predict that the peak in new cases in the USA would have occurred near March 24 if the intervention had begun on March 10. Instead, a 6-day delay in intervention shifted the initial peak to April 11, with a much higher value (Figure 3B).

Epidemic acceleration, obtained by differentiating Equation 13 twice, and then applying the definition of the RCO,

$$\frac{d^2N(t)}{d^2t} = \left(e^{RCO(t)} - \frac{K_T}{R_0}\right)\frac{dN(t)}{dt}, \tag{S2-1}$$

is calculable at any time. While its value depends on the social containment actions in effect, what is perhaps less apparent is that two countries with identical new case counts on a given day with different accelerations, will exhibit different dynamics immediately after that day.

For example, South Korea and New Zealand (Figure 2A and F) had similar new case counts (204 cases in South Korea on February 21, and 205 in New Zealand on March 25) when each imposed similar containment measures ($\frac{K_T}{R_0}$ = 0.27 in South Korea and 0.20 in New Zealand;

see Table 2). However, as evidenced by its higher RCO at the time of intervention, South Korea had a higher acceleration on February 21 compared to New Zealand on March 25. Consequently, the later number of cases in South Korea was higher than in New Zealand (Figure 2A and F).

Equation 8 clearly illustrates these lessons. Strengthening social distancing lowers $R_{eff}(t)$ by reducing R_0 and, thereby, the epidemic slows. Earlier and stronger interventions in countries with indigenously elevated levels of social interaction, are necessary to stop an epidemic in the first stages. Reopening, enacted too early, reignites an epidemic, considerably increasing the number of cases. The immense difference in the size of the effects, driven by a few days of delay, derives from the doubly exponential nature of the underlying relationships.

Supplement 3. Controlling and Ending an Ongoing Epidemic

If the situation warrants, the time needed to reach a desired reduction in $\frac{dN(t)}{dt}$, at a future time, $t + t_{target}$, can be calculated for a given level of social interaction by first defining the desired fractional reduction in cases as D_{tf},

$$D_{tf} = \frac{\frac{dN(t+t_{target})}{dt}}{\frac{dN(t)}{dt}} = \frac{New\ Case\ Target\ Rate}{Current\ New\ Case\ Rate}. \tag{S3-1}$$

Using the derivative of Equation 13, the following expression emerges from Equation S3-1,

$$D_{tf} = e^{-R_0\,(e^{-\frac{K_T}{R_0}(t+t_{target})} - e^{-\frac{K_T}{R_0}(t)})} e^{-\frac{K_T}{R_0}t_{\text{target}}}. \tag{S3-2}$$

If $t \gg t_{target}$, then $e^{-\frac{K_T}{R_0}(t+t_{target})} - e^{-\frac{K_T}{R_0}(t)} \approx 0$, and the following can be deduced from Equation S3-2,

$$t_{target} = -\frac{R_0 \ln(D_{tf})}{K_T}. \tag{S3-3}$$

Equation S3-3 quantitates the number of days, t_{target}, that a level of social containment, R_0, must continue to reach a desired reduction of the current daily cases.

We can use Equation S3-3 to predict the number of days the populace must sustain a given level of intervention, $\frac{K_T}{R_0}$, to reduce the new daily cases by a target fraction, D_{tf}. Thus, using Equation S3-3, we see that a country targeting a 90% reduction of new cases per day (e.g., from 50,000 to 5,000 cases per day, $D_{tf} = 0.1$), can attain its target in about 12 days by imposing a containment level of $\frac{K_T}{R_0} = 0.2$.

As an example of this scenario, the South Korea and New Zealand data show that both countries achieved a value of $\frac{K_T}{R_0}$ close to 0.2 for the time necessary to produce a 90% reduction. In agreement with the KMES, it took 13 days for South Korea (March 3–16) and 15 days for New Zealand (April 2–15) to reduce their new cases by 90% between the dates shown.

Returning to the planning example, after achieving the first 90% reduction, a reasonable next step might be to relax social containment to a level that allows the economy to still be viable, while preventing the epidemic from erupting again. We can again find the level of $\frac{K_T}{R_0}$

necessary to achieve a chosen target using Equation S3-3. If an additional 90% reduction in new cases per day is desired, and a period of 90 days is tolerable to achieve that reduction, then affected population must attain and maintain a new interaction level where $\frac{K_T}{R_0} \approx 0.026$. This equates to a 90-day period during which each person can on average infect ten specific people. Note that this is three times *less* stringent than the original USA shutdown level in April (Table 2). Thus, with a well-planned approach, a country can reduce its new daily cases by 99% in approximately 100 days, control the epidemic, and simultaneously maintain economic viability.

If even 0.026 is too restrictive, we can choose a still lower $\frac{K_T}{R_0}$, but it must be large enough to avoid a new outbreak. A lower bound for the new value of $\frac{K_T}{R_0}$, high enough to prevent an outbreak, can be found by setting the left-hand side of Equation S2-1 to zero, then solving for $\frac{K_T}{R_0}$:

$$e^{RCO(t)} < \frac{K_T}{R_0}. \tag{S3-4}$$

We can easily monitor the progress of interventions using the RCO (Equation 41), as the curve for South Korea illustrates (Figure 1A). Had this country maintained the initial level of distancing measures, the data would have followed the initial slope. However, the actual data departed from the slope, heralding that failures in (or relaxation of) social distancing, occurred during the indicated period (Campbell 2020) (circled data, Figure 1A). Because it summarizes epidemic dynamics, we can use the RCO to continuously determine the effectiveness of implemented measures and whether they need adjustment.

Supplement 3.1 Outbreaks

We can see from Equation 41 that if the social interactions decline (lower R_0), the slope of the RCO curve steepens, and, conversely, if the interactions increase, the slope lessens. Therefore, if the value of K_T does not change due to a change in the disease transmissibility, the RCO slope is a metric for monitoring population interactions. It is also clear that, under the assumptions used to develop the KMES, $\frac{K_T}{R_0}$ must always be greater than zero, and the RCO slope can never become positive. However, this only holds under three conditions: 1) immunity persists, 2) no new infections are introduced from outside the area, 3) any change, $\Delta R_0(t)$, due to new susceptible people entering existing N-groups or infectious people entering previously unaffected population areas must be an order of magnitude smaller than the new infections, $K_T(t)\Delta t$. We call the latter two conditions the assumption that the epidemic is contiguous.

If new infections enter a portion of the population consisting only of susceptible people, then the assumption of contiguousness does not hold. This is a common situation when infected people travel from an infected area to a previously uninfected area and cause an outbreak.

If assumptions 2 and 3 no longer hold, to predict the number of cases in an epidemic affected by an outbreak, we must modify Equation 13. If when $t = 0$, $N(0) = 1$; and introducing the notation R_{0x} where x denotes the number (order in time) of the outbreak, Equation 13 becomes:

$$N(t) = e^{-R_{01}(e^{-\frac{K_T}{R_{01}}\mathrm{t}}-1)}. \tag{S3-5}$$

If a new outbreak occurs in a previously unaffected area of a country, then we modify Equation S3-5 as follows:

$$N(t) = e^{-R_{01}(e^{-\frac{K_T}{R_{01}}\mathrm{t}}-1)} + I_2 e^{-R_{02}(e^{-\frac{K_T}{R_{02}}(\mathrm{t}-t_2)}-1)}, \tag{S3-6}$$

where I_2 is the number of infectious people who started the new outbreak, R_{02} is the replication number in the new outbreak area, and t_2 is the time the new outbreak occurs. Although we have assumed that K_T remains the same throughout this illustration, if K_T changes in a subset of the population, then the notation, K_{Tx}, can track the populations with the new transmissibility.

Equation S3-6 written in a general form is,

$$N(t) = e^{-R_{01}(e^{-\frac{K_T}{R_{01}}\mathrm{t}}-1)} + I_2 e^{-R_{02}(e^{-\frac{K_T}{R_{02}}(\mathrm{t}-t_2)}-1)} \ldots + I_{\mathrm{x}} e^{-R_{0x}(e^{-\frac{K_T}{R_{0x}}(\mathrm{t}-t_x)}-1)}, \tag{S3-7}$$

where x denotes the outbreak number and $\mathrm{t} > t_2 > t_3 > \cdots > t_x$. For each outbreak t_x, R_{0x}, and I_x need to be found independently.

While an epidemic is underway a positive slope detected in an RCO curve indicates that an outbreak has occurred, and that policy makers must take immediate action, within days, to strengthen intervention measures if they are to prevent the outbreak from overwhelming prior progress in controlling the epidemic.