



Research



Cite this article: Kollepara PK, Chisholm RH, Kiss IZ, Miller JC. 2024 Ethical dilemma arises from optimizing interventions for epidemics in heterogeneous populations. *J. R. Soc. Interface* **21**: 20230612.

<https://doi.org/10.1098/rsif.2023.0612>

Received: 20 October 2023

Accepted: 4 January 2024

Subject Category:

Life Sciences–Mathematics interface

Subject Areas:

biomathematics

Keywords:

infectious disease, intervention, trolley problem, mathematical modelling, ethics

Author for correspondence:

Joel C. Miller

e-mail: joel.c.miller.research@gmail.com

Electronic supplementary material is available online at <https://doi.org/10.6084/m9.figshare.c.7041930>.

Ethical dilemma arises from optimizing interventions for epidemics in heterogeneous populations

Pratyush K. Kollepara¹, Rebecca H. Chisholm^{1,2}, István Z. Kiss³ and Joel C. Miller¹

¹Department of Mathematical and Physical Sciences, La Trobe University, Melbourne, Australia

²Melbourne School of Population and Global Health, The University of Melbourne, Melbourne, Australia

³Network Science Institute, Northeastern University London, London E1W 1LP, UK

PKK, 0000-0002-7794-7759; RHC, 0000-0002-7830-793X; IZK, 0000-0003-1473-6644; JCM, 0000-0003-4426-0405

Interventions to mitigate the spread of infectious diseases, while succeeding in their goal, have economic and social costs associated with them. These limit the duration and intensity of the interventions. We study a class of interventions which reduce the reproduction number and find the optimal strength of the intervention which minimizes the final epidemic size for an immunity inducing infection. The intervention works by eliminating the overshoot part of an epidemic, and avoids a second wave of infections. We extend the framework by considering a heterogeneous population and find that the optimal intervention can pose an ethical dilemma for decision and policymakers. This ethical dilemma is shown to be analogous to the trolley problem. We apply this optimization strategy to real-world contact data and case fatality rates from three pandemics to underline the importance of this ethical dilemma in real-world scenarios.

1. Introduction

Infectious disease epidemics have been suppressed and mitigated using a combination of non-pharmaceutical interventions (NPIs) such as lock downs, social distancing, mask wearing and contact tracing, and by pharmaceutical interventions such as immunizing the population using vaccines. In the absence of vaccines, NPIs are the primary option. However, NPIs, and in particular, lock downs, can have significant economic, mental health and social costs associated with them. Instead of protracted or repeated lock downs (as observed during the COVID-19 pandemic), a one-shot intervention has been suggested as a possible alternative for diseases that induce immunity upon recovery from infection. An intense but short-duration lockdown is imposed near the peak of the epidemic to stop the transmission during the overshoot phase of the epidemic and reduce the final size (total number of infections) to the herd immunity threshold of the epidemic (number immune in the population required to stop the growth in infections) [1]. The overshoot phase is when the number of active infections start to decline (effective reproduction number is less than one), but a significant number of new infections are created. The overshoot is the difference between the final size and herd immunity threshold. Thus such interventions reduce the overshoot to zero (see Glossary in the electronic supplementary material for detailed definitions of technical terms).

In this work, we explore an alternative strategy to achieve the same outcome through a prolonged but weaker intervention instead of a short and intense intervention. Such an intervention, if implemented early, will have the added benefit of reducing and delaying the peak of the epidemic as well, in contrast to the one-shot intervention [1]. As with the one-shot intervention, the rationale of this strategy is to calibrate the intervention in such a manner that the final size of the

mitigated epidemic is identical to the herd immunity threshold of the original epidemic. Therefore, when the intervention ends, there is no risk of further introductions developing into future epidemics or a second wave of infections. We show that this strategy is an optimal strategy for minimizing the final size in the long term.

In the context of COVID-19 modelling, research on optimal interventions has attempted to include economic costs along with the objective of reducing infections: using detailed agent-based models [2] and fine-tuned intervention strategies [3,4], a balance is sought between socio-economic and health costs to minimize the total cost [5], or the claim that interventions reduce the economic well-being of a society has been challenged [6,7]. Optimal interventions have also been studied as resource allocation problems where a limited stockpile of vaccine is available or a limited 'amount' of social distancing is acceptable and the objective is to find the distribution of the intervention that minimizes the reproduction number or a health-related objective function [8–10].

We do not include economic costs in an explicit manner in our model. The amount of reduction in \mathcal{R}_0 can be interpreted as the cost—the higher the reduction in \mathcal{R}_0 , higher the social and economic cost of intervention. The calculations involved in finding the optimal strategy mainly rely on the knowledge of the basic reproduction number (or the next generation matrix). We show that in populations with transmission heterogeneity, implementing an optimal intervention to minimize the final size could involve a moral/ethical dilemma for decision-makers, which is analogous to the commonly known trolley problem [11,12]. The dilemma arises as a result of transmission heterogeneity in the population. We performed a literature search with relevant keywords and were unable to find any research that examined NPIs with an ethical dilemma (see electronic supplementary material for keywords). A pre-print, Ragonnet *et al.* [13], had a similar approach in that they optimized synthetic contact matrices from various European countries to minimize deaths or years of life lost by achieving herd immunity for the COVID-19 epidemics. Their model and intervention scenarios are quite complex: a transmission model with six stages of infection, waning immunity and duration of intervention. They find through numerical methods that increasing transmission in younger age groups is required to minimize the years of life lost for the COVID-19 epidemics. Another article, Babajanyan & Cheong [14], also used this strategy of achieving herd immunity in a model with susceptible \rightarrow infected \rightarrow recovered (SIR) disease and resource growth dynamics in the context of COVID-19. This study did not explore any strategies where transmission is increased. Our work differs from the above-mentioned studies as we explore strategies that increase transmission, in detail and discuss the ethical dilemma. The model we use has the minimal complexity required to explore the underlying mechanisms of this ethical dilemma for a wide range of basic reproduction numbers and to show the impact of heterogeneity and population structure on epidemics and interventions.

In the following sections, we explain the modelling framework, results of our analysis, and conclude with a discussion of our modelling assumptions and the ethical dilemma that decision-makers could face.

2. Methods

We use deterministic SIR and SIR-like models to study the optimal intervention. In §§2.1 and 2.2, we explain the

models used for a homogeneous population and for a heterogeneous population, respectively, in addition to describing the calculations for finding the optimal intervention. In §2.3, we describe how this optimization strategy is applied to real-world data and in §2.4, we explain how an optimal intervention can be found if there is a delay in the start of the intervention.

2.1. Homogeneous population

We use an SIR model with the variables s , i and r to represent the fractions of individuals in the total population who are susceptible, infected and recovered, respectively [15,16]. The population is assumed to be closed (no entry/exit) and it is normalized such that $s + i + r = 1$.

In this case, the final size of the epidemic is completely determined by the basic reproduction number \mathcal{R}_0 and can be obtained using the following equation [15,16]:

$$\ln \frac{s(t_2)}{s(t_1)} = -\mathcal{R}_0(r(t_2) - r(t_1)), \quad (2.1)$$

where $s(t_1)$ and $s(t_2)$ are the fractions of susceptible and $r(t_1)$ and $r(t_2)$ are the fractions of recovered individuals in the population at time instants, t_1 and t_2 . Using the conditions $i(t_1) \approx 0$, $r(t_1) = 0$ and $i(t_2) \approx 0$, which describe the population at the start and end of an epidemic, the well-known final size relation can be obtained [15–18]

$$r(\infty) = 1 - e^{-\mathcal{R}_0 r(\infty)}. \quad (2.2)$$

An intervention that reduces transmission would affect the basic reproduction number as $\mathcal{R}_0 \rightarrow \mathcal{R}_0(1 - c)$ where $0 \leq c \leq 1$. In the case of a homogeneous population, herd immunity is achieved when the fraction of susceptible individuals in the population is less than $1/\mathcal{R}_0$. Therefore, we substitute $s(t_2) = 1/\mathcal{R}_0$ and $s(t_1) = 1$ and solve for c . We find the optimal reduction in the basic reproduction number is

$$c = 1 - \frac{\ln \mathcal{R}_0}{\mathcal{R}_0 - 1}. \quad (2.3)$$

We verify this analytical result in §3.1 by simulating an epidemic where \mathcal{R}_0 is changed to $\mathcal{R}_0(1 - c)$ in the early stage of the epidemic and the intervention is switched off once the active infections, $i(t)$, decline to a negligible number.

2.2. Heterogeneous population

In a heterogeneous population, individuals may be further stratified into groups. To represent the fraction of individuals in the total population who belong to a group k , we use the variables s_k , i_k and r_k such that $s_k + i_k + r_k = n_k$, where n_k is the proportion of the population who belong to group k and $\sum_k n_k = 1$. Heterogeneity in transmission characteristics can affect the behaviour of epidemics in a significant manner. Epidemics in populations with different transmission structures but identical reproduction numbers can have widely different final sizes. An epidemic in a heterogeneous population can be described by the following SIR-like model, assuming identical duration of infection for all groups and measuring time in the units of the average infection duration,

$$\frac{ds_k}{dt} = -s_k \sum_l B_{kl} i_l, \quad (2.4)$$

$$\frac{di_k}{dt} = \left(s_k \sum_l B_{kl} i_l \right) - i_k \quad (2.5)$$

and
$$\frac{dr_k}{dt} = i_k. \quad (2.6)$$

Table 1. The age groups used in the contact matrix from [20], their names used in this article, size of the group (as a proportion of the total population), approximate estimates of case fatality rates (CFRs) obtained from [21–23].

age group (years)	group name	group size (Netherlands)	CFR (COVID-19/multinational)	CFR (2009 flu/Mexico)	CFR (1918 flu/USA)
1–5	children	0.06	0.000031	0.0096	0.0168
6–12	pre-adolescents	0.09	0.000008	0.0038	0.0047
13–19	adolescents	0.11	0.000028	0.004	0.0082
20–39	young adults	0.34	0.000214	0.026	0.0272
40–59	middle adults	0.23	0.001807	0.054	0.0149
60+	elderly	0.17	0.029520	0.059	0.0378

The term B_{kl} is the average number of infectious contacts that an individual in group l causes in group k . The next generation matrix, \mathbf{G} [19], can be constructed for this system with entries

$$G_{kl}(t) = s_k(t)B_{kl}. \quad (2.7)$$

The term G_{kl} is the expected number of infections that would be caused in fully susceptible group k by an infected individual in group l . The dominant eigenvalue of \mathbf{G} gives the reproduction number of the system [19].

The epidemic sizes for this model are given in [17,18]

$$r_k(\infty) = n_k(1 - e^{-\sum_l B_{kl}r_l(\infty)}). \quad (2.8)$$

It should be noted that we normalize the final size for the heterogeneous population to be the number of infections in a group as a fraction of the total population.

The recipe for optimization is similar to the homogeneous case. In the heterogeneous case, we find the level set where the reproduction number is equal to one (analytically in the case of two groups and numerically for more groups), which is the infinite set of values of s_k that would achieve the herd immunity threshold. Then, we optimize subject to this constraint to find the values, s_k^* , that minimizes the cost function (the final size or a weighted sum of final sizes of each group). From this, we obtain $r_k^* = n_k - s_k^*$. Finding the level set requires finding the proportion of susceptibles of each group, s_k , which would ensure that the reproduction number (top eigenvalue of \mathbf{G}) is equal to one. In equation (2.8), the final sizes, $r_k(\infty)$, can be replaced by the optimal final sizes, r_k^* , to find the optimal contact matrix. Comparing the original contact matrix with the optimal one tells us how the contact structure of a population must be changed in order to obtain the optimal outcome.

A crucial point to note here is that unlike with the homogeneous population, it is possible for certain elements of \mathbf{B} and for certain final sizes to increase, $r_k^* > r_k(\infty)$, in order to minimize the cost function. In other words, the optimal intervention corresponds to an increase in transmission within certain groups or among pairs of groups. In such cases, the change in reproduction number cannot be a measure of the economic or social cost. Nonetheless, this leads to some interesting results which are presented in the next section.

A weighted cost function which is a weighted sum of the final sizes in each group is useful when we are interested in minimizing a certain outcome of infections rather than the number of infections, for example, deaths or hospitalizations. The optimization problem of finding the state of the population which minimizes a general objective function and fulfils the herd immunity condition can be solved semi-analytically for the case of two groups and is presented in the electronic supplementary material. For more than two groups, we solve the optimization problem numerically.

A schematic diagram of the optimization procedure for both homogeneous and heterogeneous populations is shown in electronic supplementary material, figure S6.

2.3. Real-world contact matrix

We used a contact matrix calculated using surveys from a sample population stratified into six age groups in the Netherlands [20]. The contact matrix scaled by a disease-specific parameter gives the next generation matrix. Using the next generation matrix and the age distribution, the optimal intervention for a given cost function can be obtained. We calculated the optimal intervention using this contact matrix for a range of \mathcal{R}_0 values and four different cost function weightings (an unbiased cost function and three from observed case fatality rates (CFRs) of 2009 pandemic in Mexico, 1918 pandemic in the USA and COVID-19 pandemic) [21–23]. The age groups, their population sizes and CFRs are shown in table 1. Note that the age stratification used in the CFR study for the 1918 pandemic do not match exactly with the age groups of the contact matrix and furthermore the estimates were extracted from figures. For the 2009 pandemic and COVID-19, CFR was reported with a high age resolution but the size of the age groups was not immediately available. Due to lack of data on infection fatality rates, we are using the CFRs as a proxy for the probability that an infected individual dies. Thus the CFR values and the results relying on them are meant to be for illustration purposes only.

The severity of the dilemma in the optimal intervention can be quantified through the number of infections (or deaths) caused due to the intervention per infection (or death) prevented. It can be calculated using

$$\begin{aligned} &\text{severity of dilemma (for infections)} \\ &= \frac{\text{sum of all increases in final sizes}}{\text{sum of all the decreases in final size}} \quad (2.9) \end{aligned}$$

and

$$\begin{aligned} &\text{severity of dilemma (for deaths)} \\ &= \frac{\text{sum of all increases in deaths}}{\text{sum of all decreases in deaths}}. \quad (2.10) \end{aligned}$$

2.4. Delayed intervention in a homogeneous population

In the above sections, we have assumed that the basic reproduction number (or the next generation matrix) is a known entity and therefore an intervention is implemented right at the start of the epidemic. Calculating the strength of the optimal intervention requires knowledge of the reproduction number, the intervention would have to start after the epidemic has been established and enough observational data have been

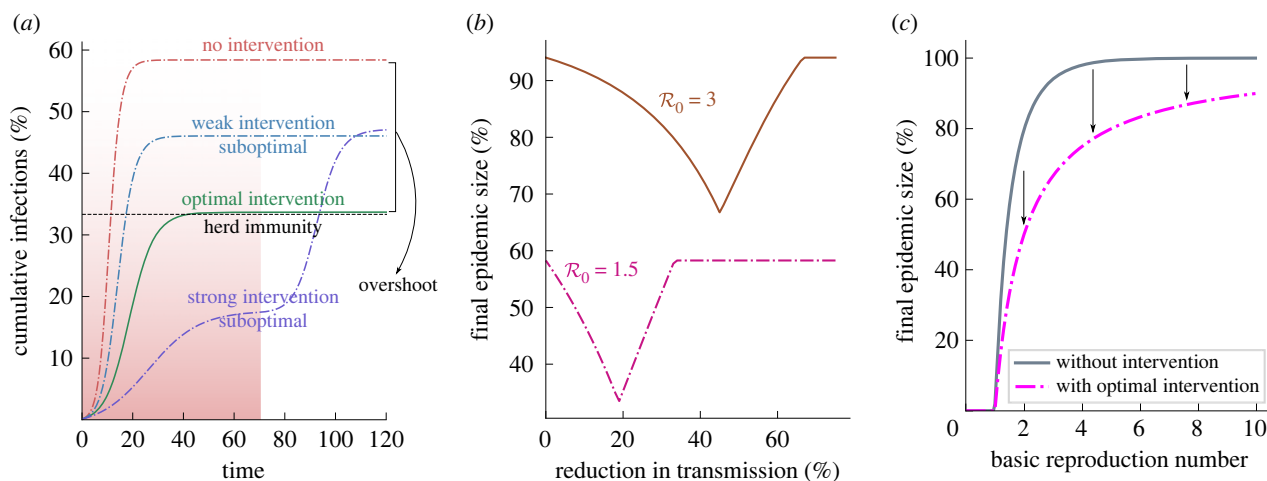


Figure 1. (a) Simulation of four types of intervention for an epidemic with $\mathcal{R}_0 = 1.5$ in a homogeneous population: (i) no intervention—leads to largest final size, (ii) weak intervention—reduces the final size, (iii) strong intervention—reduces final size during the intervention, but leads to a resurgence in infections once the intervention is removed, (iv) moderate intervention but optimal—final size during the intervention is same as the herd immunity threshold. (b) The global minimum for the final size shows that an optimal intervention strength exists. The resurgence of infections under a strong sub-optimal intervention is subject to certain assumptions which are discussed in the text. (c) The final size without any intervention (equation (2.2)) and the final size with optimal intervention (same as the herd immunity threshold for recovered state) is shown against the basic reproduction number.

collected to calculate the reproduction number. While the basic principle would still hold, a delay could change the strength of the optimal intervention. To find the optimal strength for a delayed intervention, we use the final size relation with the final state $s = 1/\mathcal{R}_0$, $i = 0$ and an initial arbitrary state s_L , i_L at a time instant t_L when the intervention begins. We replace basic reproduction number in equation (2.1) with $\mathcal{R}_0(1 - c)$ and solve for

$$c = 1 - \frac{\ln s_L \mathcal{R}_0}{\mathcal{R}_0 \left(s_L + i_L - \frac{1}{\mathcal{R}_0} \right)}. \quad (2.11)$$

Using a numerical solution of the SIR equations, s_L and i_L can be found and the above equation can be solved for c . Equation (2.11) reduces to equation (2.3) when $s_L = 1$, $i_L = 0$ and $c = 1$ when $s_L = 1/\mathcal{R}_0$ and $i_L > 0$. If $s_L < 1/\mathcal{R}_0$, then $c > 1$ which is biologically meaningless and reflects the fact that the population is already below the herd immunity threshold.

2.5. Model assumptions

The homogeneous model assumes that all individuals in the population are identical and every individual is equally likely to come in contact with every other individual. To introduce some complexity in this model, we use the heterogeneous model where individuals are stratified into homogeneous groups. We are using deterministic differential equation models with continuous variables to simulate the dynamics. This means that the number of active infections can decay exponentially but can never reach zero. Thus, the models used here cannot simulate a scenario in which an intervention eliminates a disease before reaching herd immunity threshold, as was the case in Australia, New Zealand, Hong Kong, mainland China, Singapore and several other jurisdictions (broadly known as the zero-COVID strategy). Throughout the paper, we use the SIR disease progression. Therefore, our analysis applies to diseases that induce long-term immunity or for which re-infection is not possible.

3. Results

3.1. Homogeneous population

Simulation of the SIR model differential equations confirms our assertion in equation (2.3). As shown in figure 1, a

‘weak’ intervention reduces the final size but does not reduce the overshoot to zero. A strong intervention, on the other hand, reduces the final size during the intervention but a resurgence occurs as soon as the intervention ends. The final health outcome under the strong intervention is worse than (or at least comparable to) the weak intervention, while incurring a higher social and economic cost during the intervention. The resurgence occurs because the small number of infections and sufficient number of susceptibles remaining in the population lead to new infections after the intervention is lifted. An intervention that is strong enough to minimize the final number of infections, while avoiding a resurgence, is the one whose final size (during the intervention) matches the herd immunity threshold of the unmitigated epidemic. This is the optimal intervention.

3.2. Heterogeneous population

Introducing heterogeneity in the model opens up a space of interventions that is not seen in the homogeneous case. In the homogeneous case, the herd immunity threshold is defined by a single point, but in the case of a structured population, the threshold is given by a collection of points. This can be seen by considering the following: the condition required for reaching herd immunity is that the typical infected individual must not infect more than one individual. In the homogeneous case, one can randomly choose a sufficient number of individuals and immunize them to ensure that the number of infectious contacts is less than one. If the population is structured, the typical infected individual must not infect more than one individual, on an average. As long as the average number of infectious contacts is less than one, herd immunity is achieved, irrespective of how the immunization has been distributed among the various groups in the population. Thus there are infinitely many interventions that lead to herd immunity and prevent resurgence. Out of all these possibilities, we define the optimal intervention to be the one that minimizes the final size.

When the population can be described using two sub-populations or groups, the optimal intervention belongs to one of the following types: the first group is fully infected,

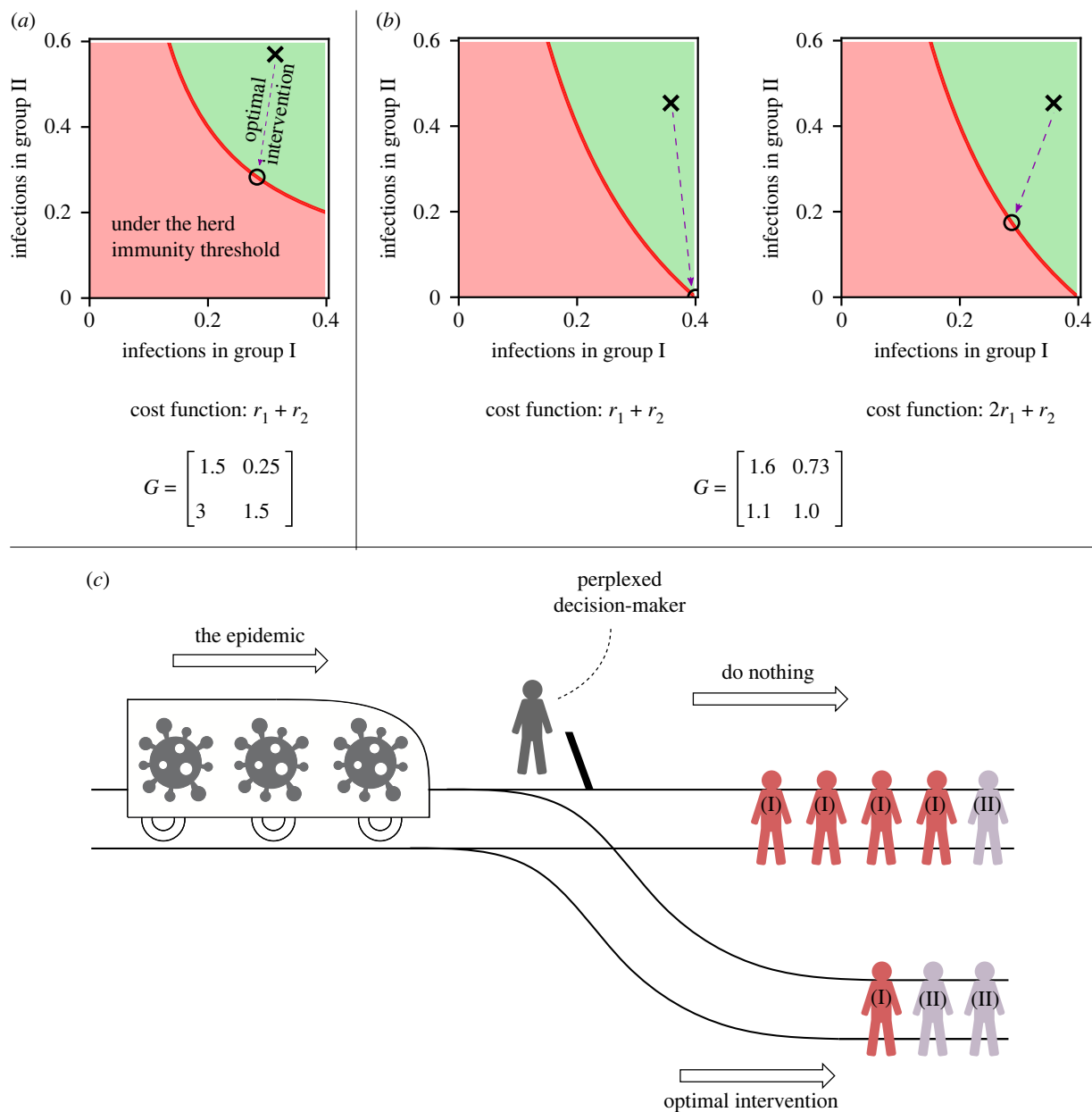


Figure 2. Population structure and nonlinearity of infectious disease dynamics lead to an ethical dilemma. (a) A plot of the infections in the two groups of the population. The corresponding cost function and next generation matrix are shown. The red curve shows the herd immunity threshold, the cross shows the final size without an intervention, the circle shows the final size if the optimal intervention is used. (b) A comparison of the interventions when the cost functions are different. The first plot shows that the optimal intervention leads to an increase in the number of infections in the first group. It is an example of the ethical dilemma of implementing the optimal intervention, which is explained further in (c). The second figure shows the plot for the cost function when the first group is given twice the weight as the second group, which means prevention of infections in the first group takes precedence over the second group. In this case, the intervention reduces the infections in both the groups. (c) The ethical dilemma involved in implementing optimal interventions is analogous to the well-known ‘trolley problem’. If the decision-maker does not act, the incoming epidemic (represented as a trolley) is going to cause many infections. If the decision-maker implements an optimal intervention (switches the tracks), the number of infections in the total population is minimized, but someone who otherwise would have been safe from infection, becomes infected (shown by the increased number of infections of group (II)).

the second group is fully infected, or none of the groups are fully infected. This creates the possibility that under the optimal intervention, the number of infections in one of the groups is larger than what would have occurred in the unmitigated epidemic, subject to the structure of the population. In figure 2, we show an example in which this occurs. This leads to an ethical dilemma wherein a certain group in the population incurs a higher cost (due to an increased number of infections) than would have happened without the intervention in order to minimize the cost for the whole population. Thus, the nonlinearity of the infectious disease dynamics, combined with population structure, lead to an ethical

dilemma for policy/decision-makers which is analogous to the well-known trolley problem [11,12] (figure 2). The trolley problem involves a set-up in which a train is going to hit a group of people who are lying on the tracks. The train cannot be stopped, but a lever can be pulled to switch the train onto a different track on which fewer people are lying. The dilemma that is posed by this situation is whether it is ethical to save more lives by ending a lesser number of different lives?

Diseases often lead to a worse health outcome (mortality rate, hospitalization rate, chance of leading to chronic conditions, etc.) in certain groups of the population (the elderly

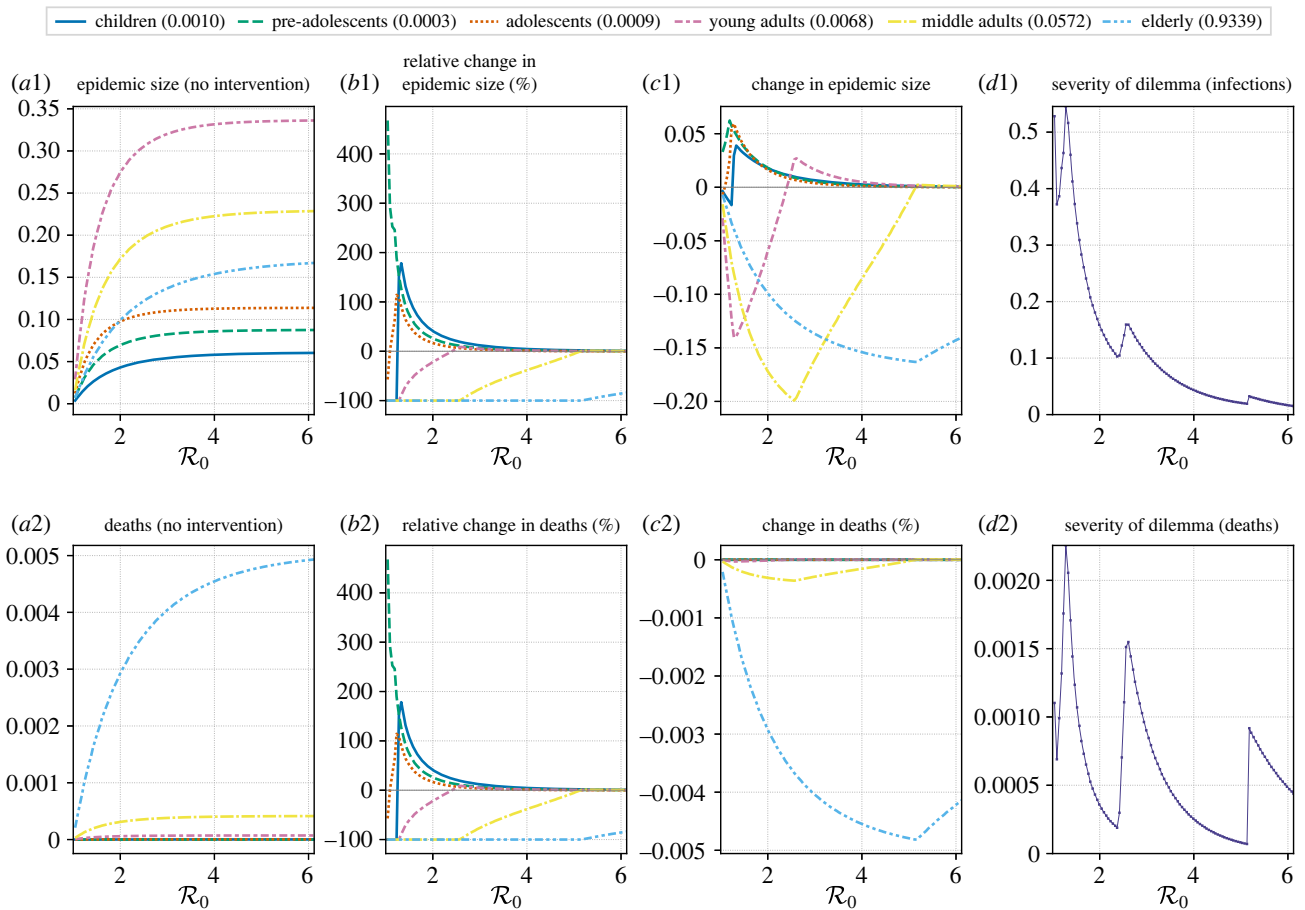


Figure 3. COVID-19 pandemic: a real-world contact matrix from a sample of the Dutch population is used to determine the effect of optimal intervention on different age groups for a range of \mathcal{R}_0 values. Estimates of case fatality rates for the COVID-19 pandemic in Mexico have been used to weight the cost function (for minimizing total deaths in the population) [23]. Rows (1) and (2) show the plots for infections and deaths, respectively. Column (a) shows the epidemic size and deaths if no intervention was performed. Column (b) shows the relative change in epidemic size and deaths under optimal intervention. Column (c) shows the magnitude of change in epidemic size and deaths under optimal intervention. Column (d) shows the severity of the ethical dilemma (see main text for definition) with \mathcal{R}_0 . The legend for columns (a–c) shows age groups and the number in parentheses shows the weight assigned to it in the cost function. These weights are proportional to the case fatality rates.

age groups for instance). Instead of minimizing the final size of the epidemic (which is the sum of final sizes in each group), it may be more prudent to minimize a cost function, which is a linear combination of the final sizes in the groups, such that a group with a worse outcome of infection is given a higher weight in the cost function. As changing the cost function would change the optimal solution, the cost function plays a role in determining the ethical dilemma. For the example shown in figure 2, the ethical dilemma is no longer present under the given weighted cost function; the infections are reduced in both groups.

3.3. Real-world contact matrix

When the basic reproduction number is close to one, at least one of the age groups is required to endure a higher final size for all the cost functions we used (figure 3; electronic supplementary material, figures S3–5). The cost functions are weighted using estimates of CFRs of the 2009 influenza pandemic, the 1918 influenza pandemic, and COVID-19 pandemic [21–23]. In addition to looking at the final sizes and how they change in various age groups, we can use the CFRs to estimate the deaths in each of the age groups and how they change with the optimal intervention.

For the COVID-19 pandemic, pre-adolescents have the lowest CFR, and it increases for higher age groups (figure 3 and table 1). Figure 3b1,b2,c1,c2 shows that as \mathcal{R}_0 is increased, the age groups start to experience an increase in infections (relative to no intervention case) in the following order—pre-adolescents, adolescents, children and finally young adults, which is also the order in which the CFR increases. Thus, the CFR may explain the nature of the ethical dilemma. At $\mathcal{R}_0 \approx 1.3$, the severity of the ethical dilemma is highest, with 0.55 new infections for every infection prevented and 0.0025 new deaths for every death prevented (figure 3d1, d2). The severity for deaths is quite low compared to other pandemics because of the large disparity in the CFR across age for COVID-19.

For the 2009 flu pandemic (electronic supplementary material, figure S4), pre-adolescents have the lowest CFR, and it increases with age (electronic supplementary material, figure S4 and table 1). Electronic supplementary material, figure S4 (columns B and C) shows that as \mathcal{R}_0 is increased, the age groups start to experience an increase in infections (relative to no intervention case) in the following order—adolescents, pre-adolescents and finally young adults, which is not in the increasing order of CFRs. Thus, the CFR does not explain the nature of the ethical dilemma. For infections, the severity of dilemma is highest at about $\mathcal{R}_0 = 1.3$,

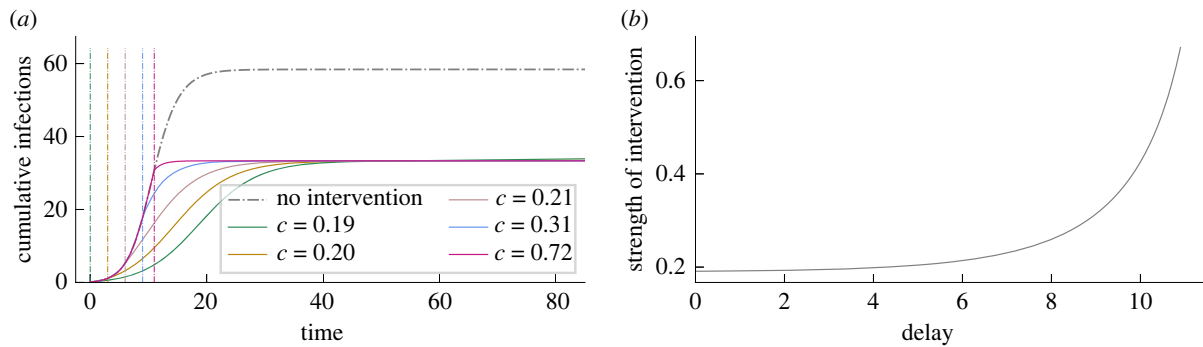


Figure 4. (a) The time series of cumulative infections ($i + r$) for optimally controlled epidemics for a range of delays in the intervention. The vertical dash-dot lines show the time at which the intervention starts, and its corresponding time series is represented by the same colour. (b) The strength of optimal intervention, c , plotted against the delay in intervention t_l using equation (2.11). The super-linear increase in the strength, c , shows the need for an early implementation. Parameters: homogeneous SIR model with basic reproduction number $\mathcal{R}_0 = 1.5$ and $\gamma = 1$. Intervention is implemented for a duration of 50 time units.

where 0.44 new infections are created for every infection prevented. For deaths, the dilemma is the most severe at both $\mathcal{R}_0 = 1.3$ and 2.25, where 0.045 new deaths are caused for every death prevented.

For the 1918 flu, the CFR with age is often described as a ‘W’ shaped curve (electronic supplementary material, figure S5 and table 1). Electronic supplementary material, figure 5 (columns B and C) shows that as \mathcal{R}_0 is increased, the age groups start to experience an increase in infections (relative to no intervention case) in the following order—adolescents, pre-adolescents, and finally middle adults, which is also the order in which the CFR increases. Thus, the CFR may explain the nature of the ethical dilemma. At the peak of severity ($\mathcal{R}_0 \approx 1.3$), 0.47 new infections are created for each infection prevented and 0.12 new deaths for every death prevented.

For realistic CFRs, the dilemma in terms of infections is quite severe at its worst, with almost one person getting infected for protecting two individuals from getting infected. Some features of the ethical dilemma are common to all three pandemics. As \mathcal{R}_0 is increased, the severity of dilemma never quite reaches zero but seems to approach zero in a non-monotonic manner and pre-adolescents and adolescents always endure an increase in infections. The nature of the ethical dilemma may be explained by the CFRs of the age-groups in the 1918 and COVID-19 case, but not in the case of the 2009 pandemic.

For an unbiased cost function (electronic supplementary material, figure S3), we see very different results. The severity of dilemma is zero for most of the \mathcal{R}_0 range. In the space where the dilemma does occur, only the young adults and adolescents experience an increase in final size. At the most severe ethical dilemma, 0.175 new infections are created for every infection prevented.

3.4. Delayed intervention

We calculate the optimal strength of the intervention and simulate the model to confirm the mathematical analysis in §2.4. Using equation (2.11), we observe that the strength of optimal intervention increases in a super-linear manner with the duration of delay. The results are presented in figure 4. As the population approaches the herd-immunity threshold, the strength of intervention approaches one, corresponding to the one-shot intervention [1].

4. Discussion

In this work, we have examined a strategy of optimal intervention which allows the epidemic to cause just enough infections to induce herd immunity, eliminate the overshoot, and prevent future introductions from becoming epidemics. In addition to minimizing the final size, this intervention would also slow down the growth of the epidemic and reduce the peak, which allows time to develop treatments and increase healthcare capacity. For a homogeneous population, the results are straightforward: decrease the transmission by a pre-determined amount so that the final size reaches the herd immunity threshold and no more. A sensitivity analysis of the homogeneous model and intervention strategy was performed where the optimal strength of intervention and the resulting final size were computed for both the actual value of \mathcal{R}_0 and a ‘measured’ value of \mathcal{R}_0 with four different error rates (see electronic supplementary material, figure S2). We find that in both underestimation and overestimation of \mathcal{R}_0 , the epidemic size is larger, but it is better to overestimate \mathcal{R}_0 .

In the case of heterogeneous transmission, our results indicate that the optimal strategy may require increasing infection in some of the groups and decreasing it in others, in order to minimize the final size for the whole population. This is analogous to the trolley problem, and it calls for a discussion around the ethics of subjecting certain groups to a higher rate of disease incidence, and the feasibility of this policy. If increasing transmission in certain groups is not viable either due to operational reasons or ethical considerations, herd immunity can still be achieved (and resurgence prevented) by reducing transmission in all groups. We have also explored the role of the cost function in determining the ethical dilemma by weighting the final sizes of different age groups using CFRs of the 1918, 2009 and COVID-19 pandemics and shown that the ethical dilemma happens in all three cases. Our work shows that even without an explicit consideration of economic and social costs of an intervention, there are challenging ethical questions to be answered for the first-order problem of minimizing the final size.

The optimal interventions shown for the 1918, 2009 and COVID-19 pandemics are not meant to be policy advice because the estimates of CFR were approximate and also because influenza and COVID-19 can be described by an SIR-like model only when new variants of the pathogen do not emerge. They are meant to show that the ethical dilemma

we have discussed in this paper is not merely a theoretical observation in the parameter space of the mathematical model but a possibility that one should be aware of for future epidemics and pandemics.

Ethical dilemmas in public health are well known, and there have been debates on prioritization based on age for vaccination during the COVID-19 pandemic [24]. Prioritizing one age group means that another group receives less protection, but under no circumstances does a discriminatory vaccine distribution policy increase the chances of infection in a group compared to the no vaccine scenario. Thus the dilemma presented in our paper is fundamentally different to a vaccine allocation dilemma and is equivalent to the trolley problem. We have used mathematical modelling to show that optimal interventions may require a policymaker to contend with a trolley-problem-like situation where the epidemic under an optimal intervention will infect someone who would not have been affected if there was no intervention. We found two works that use an intervention similar to ours, but they did not consider increase in transmission as a strategy for optimal outcomes [14] or did not discuss the ethical implications [13]. Therefore, we believe that our paper contributes to an important point of discussion with regards to optimality of NPIs.

In addition to the ethical dilemma shown through our modelling here, interventions that require increasing transmission prompt an ethical discussion in relation to disadvantaged groups. Cultural, economic and social conditions factor into the contact structure of any human population—a high number of contacts due to living in close spaces, a high susceptibility to infection due to pre-existing health conditions or poor access to healthcare facilities, etc. Mathematical models of epidemics can throw light on possible choices of policy and may even help us pick the ones that lead to optimal outcomes. But the decisions made by policymakers are intertwined with political will, their popularity and social attitudes. These eventually determine whether a particular intervention is favoured by a decision-making body [25,26]. Disadvantaged groups, across the world, do not exercise sufficient political power to represent their interests in decision-making bodies. In such a case, a decision-making body may find it convenient to subject a disadvantaged group to a higher final size in order to decrease the net final size for the whole population and achieve herd immunity. The intervention strategy presented here, always carries such risks with it; and representation of disadvantaged groups thus becomes essential, especially for a policy such as this one.

There are also some practical limitations to the strategy presented here. There would be a natural tendency for individuals to protect themselves from getting infected even if interventions are not in place, so asking individuals to increase their transmission may not be a feasible strategy [27]. The optimal interventions could require a group of individuals to fully isolate themselves from the rest of the population. Such interventions are difficult to implement,

as there would always be a small possibility for infections to be introduced into the isolated group [27]. If the transmission in other groups is increased, it would imply a larger chance of introduction into the isolated group.

We have assumed an SIR structure for disease progression in an individual. But, as long as the disease can be reasonably described by a model in which individuals do not become susceptible after getting infected, we would expect our results to be valid. A crucial detail that we have ignored is the stochastic and discrete nature of disease spread since it can capture the elimination behaviour of outbreaks, i.e. it can incorporate the difference between existence and absence of infections. The deterministic assumption and the use of continuous variables in our model means that after an intervention is over, the small number of infections present in the population will lead to another epidemic if herd immunity is not achieved (shown in figure 1, strong intervention). This however, is one of the possible outcomes. It is possible that the intervention completely eliminates all infections in the population, in which case a new epidemic does not result from any residual infections. However, even in this case, the population remains vulnerable to an epidemic due to lack of herd immunity. Thus, a new epidemic can occur if new infectious individuals are introduced into the population. Another possibility is that the epidemic may get established with a delay due to the stochastic dynamics. Factors around contact-tracing and surveillance capacity (to eliminate the disease) and travel restrictions (to prevent introduction of new infections) are important for the selecting the optimal policy response, in addition to the results presented here.

Ethics. Ethics approval was not required for this study.

Data accessibility. The software used to generate the figures is available from the Zenodo repository: <https://doi.org/10.5281/zenodo.10386309> [28]. Data on case fatality ratios came from: [https://www.thelancet.com/journals/laninf/article/PIIS1473-3099\(10\)70120-1/fulltext](https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(10)70120-1/fulltext); <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3291398>; <https://www.nature.com/articles/s41586-020-2918-0>; data on the contact network came from: <https://academic.oup.com/aje/article/164/10/936/162511>.

Supplementary material is available online [29].

Declaration of AI use. We have not used AI-assisted technologies in creating this article.

Authors' contributions. P.K.K.: conceptualization, formal analysis, investigation, methodology, software, visualization, writing—original draft, writing—review and editing; R.H.C.: conceptualization, investigation, methodology, project administration, supervision, writing—review and editing; I.Z.K.: conceptualization, methodology, supervision, writing—review and editing; J.C.M.: conceptualization, investigation, methodology, project administration, supervision, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by a La Trobe University Graduate Research Scholarship (LTGRS) and a La Trobe University Full Fee Research Scholarship (LTUFFS) for P.K.K., and startup funding from La Trobe University for J.C.M.

Acknowledgements. We thank Michael Meehan, Sarah Becirevic and Somya Mehra for helpful discussions.

References

1. Di Lauro F, Kiss IZ, Miller JC. 2021 Optimal timing of one-shot interventions for epidemic control. *PLoS Comput. Biol.* **17**, e1008763. (doi:10.1371/journal.pcbi.1008763)
2. Blakely T *et al.* 2021 Determining the optimal COVID-19 policy response using agent-based

- modelling linked to health and cost modelling: case study for Victoria, Australia. *medRxiv*, 2021.01.11.21249630. (doi:10.1101/2021.01.11.21249630)
3. Britton T, Leskelä L. 2023 Optimal intervention strategies for minimizing total incidence during an epidemic. *SIAM J. Appl. Math.* **83**, 354–373. (doi:10.1137/22M1504433)
 4. Cheong KH, Wen T, Lai JW. 2020 Relieving cost of epidemic by Parrondo's paradox: a COVID-19 case study. *Adv. Sci.* **7**, 2002324. (doi:10.1002/advs.202002324)
 5. Samuel J, Sinha S. 2021 Optimal control in pandemics. *Phys. Rev. E* **103**, L010301. (doi:10.1103/PhysRevE.103.L010301)
 6. Dorn F *et al.* 2022 The common interests of health protection and the economy: evidence from scenario calculations of COVID-19 containment policies. *Eur. J. Health Econ.* **24**, 67–74. (doi:10.1007/s10198-022-01452-y)
 7. Siegenfeld A, Bar-Yam Y. 2020 Minimizing economic costs for COVID-19. New England Complex Systems Institute. See <https://necsi.edu/s/EconTradeoffs2.pdf>.
 8. Wallinga J, van Boven M, Lipsitch M. 2010 Optimizing infectious disease interventions during an emerging epidemic. *Proc. Natl Acad. Sci. USA* **107**, 923–928. (doi:10.1073/pnas.0908491107)
 9. Goldstein E, Apolloni A, Lewis B, Miller JC, Macauley M, Eubank S, Lipsitch M, Wallinga J. 2010 Distribution of vaccine/antivirals and the 'least spread line' in a stratified population. *J. R. Soc. Interface* **7**, 755–764. (doi:10.1098/rsif.2009.0393)
 10. Abell IR, McCaw JM, Baker CM. 2023 Understanding the impact of disease and vaccine mechanisms on the importance of optimal vaccine allocation. *Infect. Dis. Model.* **8**, 539–550. (doi:10.1016/j.idm.2023.05.003)
 11. Foot P. 2002 *Virtues and vices*. Oxford, UK: Oxford University Press.
 12. Thomson JJ. 1976 Killing, letting die, and the trolley problem. *Monist* **59**, 204–217. (doi:10.5840/monist197659224)
 13. Ragonnet R *et al.* 2020 Optimising social mixing strategies to mitigate the impact of COVID-19 in six European countries: a mathematical modelling study. *medRxiv*, 2020.08.25.20182162. (doi:10.1101/2020.08.25.20182162)
 14. Babajanyan SG, Cheong KH. 2021 Age-structured SIR model and resource growth dynamics: a COVID-19 study. *Nonlinear Dyn.* **104**, 2853–2864. (doi:10.1007/s11071-021-06384-5)
 15. Kermack WO, McKendrick AG. 1927 A contribution to the mathematical theory of epidemics. *Proc. R. Soc. Lond. A* **115**, 700–721. (doi:10.1098/rspa.1927.0118)
 16. Hethcote HW. 2000 The mathematics of infectious diseases. *SIAM Rev.* **42**, 599–653. (doi:10.1137/S0036144500371907)
 17. Miller JC. 2012 A note on the derivation of epidemic final sizes. *Bull. Math. Biol.* **74**, 2125–2141. (doi:10.1007/s11538-012-9749-6)
 18. Andreasen V. 2011 The final size of an epidemic and its relation to the basic reproduction number. *Bull. Math. Biol.* **73**, 2305–2321. (doi:10.1007/s11538-010-9623-3)
 19. Diekmann O, Heesterbeek J, Metz J. 1990 On the definition and the computation of the basic reproduction ratio R_0 in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* **28**, 365–382. (doi:10.1007/BF00178324)
 20. Wallinga J, Teunis P, Kretzschmar M. 2006 Using data on social contacts to estimate age-specific transmission parameters for respiratory-spread infectious agents. *Am. J. Epidemiol.* **164**, 936–944. (doi:10.1093/aje/kwj317)
 21. Nishiura H. 2010 Case fatality ratio of pandemic influenza. *Lancet. Infect. Dis.* **10**, 443–444. (doi:10.1016/S1473-3099(10)70120-1)
 22. Taubenberger JK, Morens DM. 2006 1918 influenza: the mother of all pandemics. *Emerg. Infect. Dis.* **12**, 15–22. (doi:10.3201/eid1209.05-0979)
 23. O'Driscoll M, Ribeiro Dos Santos G, Wang L, Cummings DA, Azman AS, Paireau J, Fontanet A, Cauchemez S, Salje H. 2021 Age-specific mortality and immunity patterns of SARS-CoV-2. *Nature* **590**, 140–145. (doi:10.1038/s41586-020-2918-0)
 24. Lee A, Morling J. 2022 COVID-19 vaccine dilemmas. *Public Health* **202**, 10–11. (doi:10.1016/j.puhe.2021.01.009)
 25. Palus S. 2020 America never did a real lockdown. *Slate*. See <https://slate.com/technology/2020/11/america-never-did-lockdown.html>.
 26. Yerushalmy J. 2022 Zero-COVID policy: why is China still having severe lockdowns? *The Guardian*, 29 November 2022. See <https://www.theguardian.com/world/2022/nov/29/china-zero-covid-policy-what-is-it-and-why-lockdowns-quarantine-protests>.
 27. Smith CA, Yates CA, Ashby B. 2022 Critical weaknesses in shielding strategies for COVID-19. *PLoS Global Public Health* **2**, e0000298. (doi:10.1371/journal.pgph.0000298)
 28. Kollepara PK, Chisholm RH, Kiss IZ, Miller JC. 2024 Ethical dilemma arises from optimizing interventions for epidemics in heterogeneous populations. Zenodo. (doi:10.5281/zenodo.10386309)
 29. Kollepara PK, Chisholm RH, Kiss IZ, Miller JC. 2024 Ethical dilemma arises from optimizing interventions for epidemics in heterogeneous populations. Figshare. (doi:10.6084/m9.figshare.c.7041930)