

What makes for effective pandemic policy?

David Yuan, Nathan Cantafio

Abstract

In light of the COVID-19 pandemic, understanding the impact of public health interventions on epidemic spread has become crucial for managing future outbreaks. A common framework for pandemic modelling is the SIR (Susceptible/Infected/Recovered) model which uses differential equations to model disease spread. There are many extensions of this model that add more groups (such as SIVQR – Susceptible/Infected/Vaccinated/Quarantined/Recovered), and there are innumerable different policy interventions/characteristics of the disease one can consider in the model. This report aims to determine which policy measures are most important in managing disease spread.

1. Introduction

The question of interest is determining which policy levers are the most effective at managing disease spread during an outbreak. Some examples of policy levers are a quarantine policy for infected individuals, blanket social distancing measures, mask mandates, vaccine programs, etc. Inspired by the model presented in Turkyilmazoglu, [2022](#), we will simulate the evolution of a pandemic under various scenarios.

To measure how effective a particular strategy is, we would like to look at the resulting load on the healthcare system. This depends on how many people get sick, what kinds of people are getting sick (is it mostly elderly or healthy young people?), how long they are sick for. If at any point the load is too high for the healthcare system to handle, people will be left without treatment and potentially die. Thus we aim to keep the maximum load over the entire period as low as possible, i.e. we want to “flatten the curve”.

In practice, we don’t have a way of directly measuring the load. A good proxy however, is the total number of people who are sick. Which in our model includes those who are infectious and the infectious people who have been put in quarantine.

The remainder of this report is organized as follows: Section [2](#) outlines the experimental design, highlighting the rationale behind factor selection and level choices. Section [3](#) presents the results of simulations and statistical analyses. Section [4](#) delves into the implications of these results, offering insights and recommendations for policymakers. The Appendix includes more in depth discussion of the model and the simulation setup, as well as all the relevant tables and figures.

2. The experimental design

In the MATLAB code there are eight parameters of interest that we can tune. They are in no particular order: `prob_spread`, `recovery_rate`, `vac_eff`, `prob_symp`, `isolation`, `vac_rate`, `quar_dur`, `num_daily`.

These represent (respectively) the probability that an unvaccinated individual will catch the disease from a single interaction with an infectious individual, the average rate of recovery for an infected individual in days, the proportional effectiveness of the vaccine (for example 0.1 means that vaccinated individuals are 10% as likely to get infected compared to if they were unvaccinated), the probability that an infected individual shows symptoms, the social isolation policy in place (i.e. at the extreme ends: a social isolation policy of 0 corresponds to a policy that leads to no reduction in the number of daily interactions, and a social isolation policy of 1 corresponds to a policy that leads to a 90% reduction in the number of daily interactions), the rate at which the susceptible population is vaccinated in days, the duration that symptomatic individuals are quarantined in days, the average number of daily interactions that have the potential to spread the disease that an individual has per day.

The first of four of these parameters were randomized in simulation for each observation. This choice was made to reduce the number of effective runs, and because these parameters are traits of the particular disease and would be hard (or potentially impossible) to measure in the real world. More detail is given in Appendix A and Table 2. The last four of these parameters were considered as treatment/blocking factors and their levels were thus pre-designed and recorded.

The design is a full 3^4 factorial. For a summary of the experimental factors, see Table 1. These levels were chosen to be realistic so that they would provide useful information and to cover a wide range of possibilities without increasing the size of the experiment excessively. The levels of `soc.iso` were chosen to capture the three main types of policy: no policy, somewhat lax “social distancing”, and harsher lockdowns where nobody interacts with anybody except for what is absolutely necessary. The levels of `vac.rate` were chosen based on data in Our World in Data, 2024 where the highest rates range between about 1–2% of the population vaccinated per day and the lowest level 0% serves as baseline with no vaccination. The levels of `quar.dur` were chosen based on the 7 and 14 day quarantine periods having been common recommendations during the COVID-19 pandemic, again with 0 (no quarantine) serving as a baseline. The levels of `num.daily` were chosen to block for variation in population density between different regions. This is a hard number to estimate, but based on data from Del Valle et al., 2007 we think that between 15 to 45 interactions per day which would correspond to moderately suburban areas and highly dense city centres respectively is reasonable.

The response of interest is `load` which is defined as the maximum proportion of sick (infectious or quarantined) individuals over 90 days. For a full description see Appendix B.

The model equation is:

$$Y_{ijkl} = \mu + \alpha_i + \nu_j + \kappa_k + \eta_\ell + (\alpha\nu)_{ij} + (\alpha\kappa)_{ik} + (\alpha\eta)_{i\ell} + (\nu\kappa)_{jk} + (\nu\eta)_{j\ell} + (\kappa\eta)_{k\ell} \\ + (\alpha\nu\kappa)_{ijk} + (\alpha\nu\eta)_{ij\ell} + (\alpha\kappa\eta)_{ik\ell} + (\nu\kappa\eta)_{jkl} + (\alpha\nu\kappa\eta)_{ijkl} + E_{ijkl}$$

where $i, j, k, \ell \in \{1, 2, 3\}$ and a full description of the terms is contained in Table 3

3. Analysis

Try to be specific about your findings, and present numerical estimated effects, recommended levels, the estimated response at the recommended levels, etc. “The high level of temperature was better at the 5% significance level” does not summarize well: the reader would have to make some more calculations to assess whether the result is practically significant. Instead, “We are 95% confident that the effect of changing temperature from 15 to 25°C is an increase in mean (a response variable) of (a confidence interval)” tells the reader immediately whether there is a result of concern.

If interaction effects are reported you should be careful to interpret them. Numerical estimates of interaction effects are not easy to understand. Rather, refer to a table of averages or an interaction plot. Talk about how the estimated effect of one factor depends on the level of another.

Do not give excessive digits for estimates, etc. Two significant digits are usually sufficient for a standard error. Give the corresponding estimate to the same accuracy, e.g., 14.37 kg with standard error 0.54 kg, 14.4 kg with standard error 5.4 kg, or 14 kg with standard error 54 kg. Always give units of measurement for estimates and standard errors.

If a transformation has been applied you will often want to convert predictions on the transformed scale back to the original scale.

4. Conclusions

The conclusions are often essentially a summary of the important results.

Evaluate what you did. What would you do differently next time if a similar experiment were conducted?

References

- Del Valle, S., Hyman, J., Hethcote, H., & Eubank, S. (2007). Mixing patterns between age groups in social networks. *Social Networks*, 29, 539–554. <https://doi.org/10.1016/j.socnet.2007.04.005>
- Our World in Data. (2024). Daily share of the population receiving a COVID-19 vaccine dose [Accessed: 2024-12-4].

Turkyilmazoglu, M. (2022). An extended epidemic model with vaccination: Weak-immune sirvi. *Physica A: Statistical Mechanics and its Applications*, 598, 127429. <https://doi.org/https://doi.org/10.1016/j.physa.2022.127429>

Appendix

A. Discussion of simulation

The model considers five distinct groups: S , I , V , Q , and R ; which are the number of susceptible, infected, vaccinated, quarantined and recovered individuals respectively. If N is the total population, we can define $s = S/N$, $i = I/N$, $v = V/N$, $q = Q/N$, $r = R/N$ and work with these population proportions. The system of differential equations is

$$\begin{cases} ds/dt = -\beta_0 si - \nu s \\ di/dt = \beta_0 si + \beta_1 vi - \mathbb{I}_Q P_{\text{sympt}} i + (1 - \gamma_Q) \kappa q - \gamma i \\ dv/dt = \nu s - \beta_1 vi \\ dq/dt = \mathbb{I}_Q P_{\text{sympt}} i - \kappa q \\ dr/dt = \gamma i + \gamma_Q \kappa q \end{cases} \quad (1)$$

where

- β_0 and β_1 are the transmission rate ($\#$ daily interactions $\times \mathbb{P}$ [spread per interaction]) for unvaccinated and vaccinated individuals respectively
- ν is the vaccination rate of susceptible individuals
- \mathbb{I}_Q is an indicator variable equalling 1 if there is a quarantine policy and 0 if there is no quarantine policy
- P_{sympt} is the probability that an infected individual is symptomatic
- γ and γ_Q are the average rates $[1/(\text{average recovery duration})]$ of recovery for infected individuals and individuals who just came out of quarantine respectively
- κ is the average rate $[1/(\text{average quarantine duration})]$ at which we take people out of quarantine

The **MATLAB** code uses Euler's method with time steps of size $k = 0.001$ to simulate the dynamics of these groups up to $T_f = 90$ days from an initial condition of 1% infected and everyone else susceptible. Parameters mentioned to be randomized in Section 2 were done so using the built in functions `randn()`. The range of possible values are shown in Table 2. The response was measured as `load = max(i + q)`.

B. Description of data

Explain all variables...

C. Tables and figures

Factor	Levels	Symbol
Isolation	Low (0), Medium (0.5), High (1)	<code>soc.iso</code>
Vaccination rate	0 ppd, 1 ppd, 2 ppd	<code>vac.rate</code>
Quarantine duration	0 days, 7 days, 14 days	<code>quar.dur</code>
Average number of daily interactions	15 ipd, 30 ipd, 45 ipd	<code>num.daily</code>

Table 1: ppd is “percent population vaccinated per day” and ipd is “average number of daily interactions per day”

Parameter	Range	Variable name
Probability of spread per interaction	[0.05, 0.15]	<code>prob_spread</code>
Recovery rate	[7, 21] days	<code>recovery_rate</code>
Vaccine effectiveness	[40%, 60%]	<code>vac_eff</code>
Proportion symptomatic	[0, 0.6]	<code>prob_symp</code>

Table 2: The above parameters were uniformly generated from the corresponding range

Term	Description
μ	an effect common to all observations
α_i	the effect on <code>load</code> of <code>soc.iso</code> at level i
ν_j	the effect on <code>load</code> of <code>vac.rate</code> at level j
κ_k	the effect on <code>load</code> of <code>quar.dur</code> at level k
η_ℓ	the effect on <code>load</code> of <code>num.daily</code> at level ℓ
$(\alpha\nu)_{ij}$	the “interaction effect” on <code>load</code> of <code>soc.iso</code> and <code>vac.rate</code>
other interactions terms	similar
E_{ijkl}	the residual error at levels i, j, k, ℓ ; distributed as $\mathcal{N}(0, \sigma^2)$
Y_{ijkl}	the recorded value of <code>load</code> at level i, j, k, ℓ

Table 3: A description of terms in the model equation

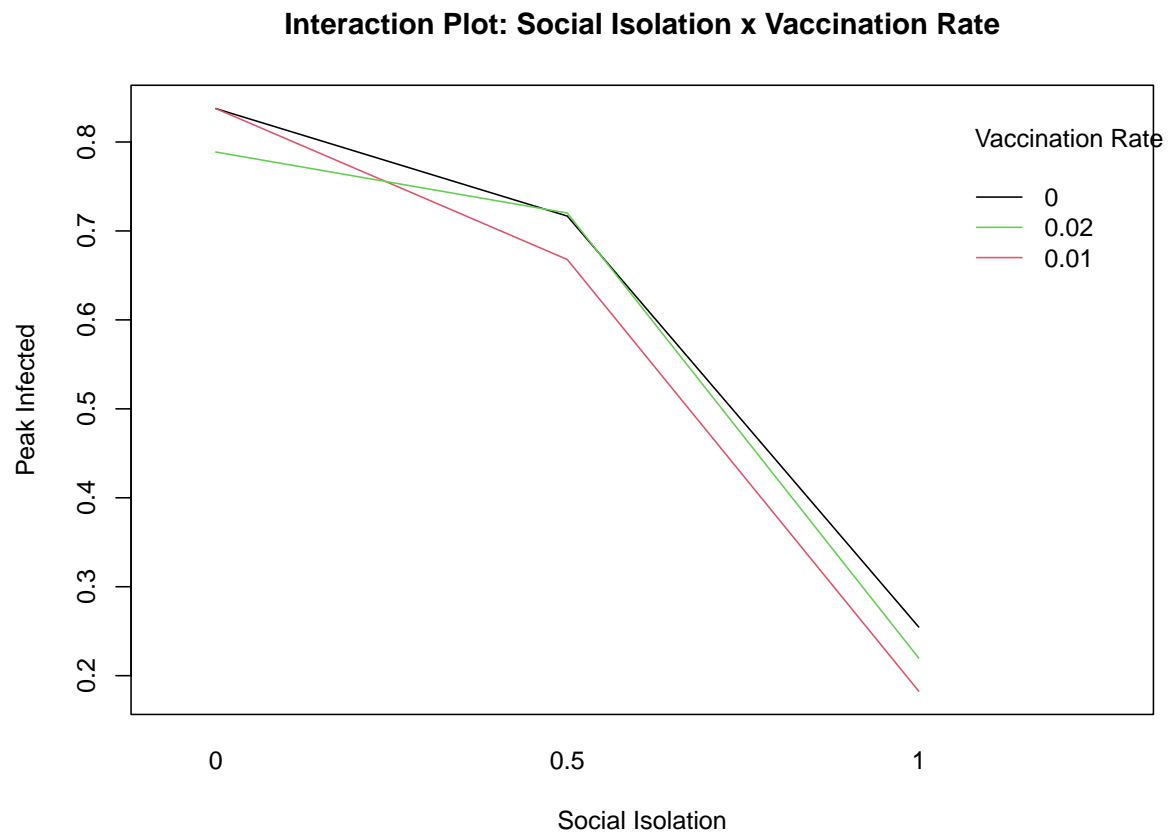


Figure 1: Interaction plot