

# SEIR Model with Mutation and Traveling

Abraham Holleran and Mike Veatch

Gordon College

December 13, 2021

## 1 Introduction

We set out to make a SEIR model that includes vaccinations and mutations and accounts for traveling between areas. It's a modified version of a cutting-edge SEIR model used to allocate vaccines. (Bertsimas, et. al, 2021) [1]. Our model uniquely attempts to describe the ebb and flow of the virus on a global scale. This allows for comparisons of different countries with different vaccination rates and different travel (flight) restrictions.

It is then embedded in an optimization model that considers the decision of how to allocate vaccines to different areas globally.

### State Variables:

The SEIR model is a deterministic system of differential equations that represents people travelling through states. There are 8 states for people in each area  $a$ , and each state is a function with respect to time.

$S_a$  Susceptible

$M_a$  Immune due to vaccination

$E_a$  Exposed

$I_a$  Infectious

$Q_a$  Self-Quarantined

$H_a$  Hospitalized

$D_a$  Dead

$R_a$  Recovered

The total number of people,  $N_a$ , is equal to the sum of the above states. See Figure 1 for inter-state interactions.

### Parameters

Each rate is in units of proportion per day.

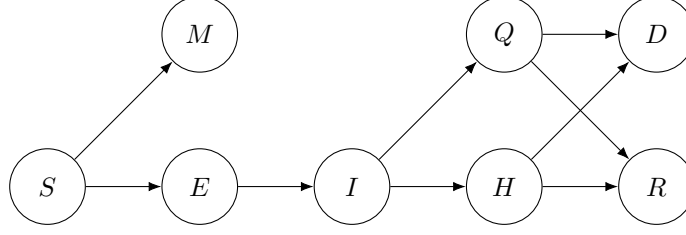


Figure 1: The states for an area  $a$

$r^d$  rate of detection, either from a self-diagnosis or from a positive test.

$p^d$  Proportion of true cases detected. Some are undetected due to limited testing, asymptomatic carriers, and detection errors.

$p^H$  Proportion of detected people who need hospitalization.

$r^R$  the rate of recovery or death out of states Q and H.

$\beta$  Vaccine effectiveness. Proportion who become immune via vaccine

$A$  The set of areas.

$T_D$  Time for a variant to dominate (days).

$m_a$  Constant mortality rate in each area.

$N_a$  Initial population in area  $a$ .

$\lambda$  Increase to infection rate of next variant due to mutation. Unitless.

$T$  Length of time horizon for model (days).

$N$  Number of infected person-days, globally, between variants.

$t_N$  Day when global infection-days exceeds  $N$ .

Defined as the first day where  $\sum_{a=1}^{|A|} \int_0^{t_N} I_a(t) dt \geq N$

$p$  proportion of people in state I that have the new variant when it is introduced in an area.

$S_a(0), E_a(0), I_a(0)$  initial population in a state

## 2 Other Variables

$V_a(t)$  The number of people vaccinated per day in each area  $a$  at time  $t$ . A decision variable.

$\alpha_a(t)$  The infection rate in area  $a$  at time  $t$ .

$\gamma_a(t)$  Governmental and societal response (unitless). Measures reduction of virus transmission.

### 3 SEIR Model

The SEIR model describes the movement of individuals through states. In this section, there are separate dynamics for each area  $a$  (region of the world). There are no age groups or risk groups. Areas interact through variants, which occur based on global infections. The Appendix 7.1 adds inter-area travel.

## 4 Model Assumptions

### 4.1 Vaccination

A proportion  $\beta$  of vaccinated individuals become immediately, permanently immune to the virus. This makes  $M$  a terminal state. See Appendix 7.3 for a version of vaccination that accounts for partial immunity.

### 4.2 Infection

People in state  $I$  stay infectious until something causes them to leave. We call this process "detection" even though there are a few ways to stop being infectious. If an infectious person gets a positive test and shows severe symptoms, they get hospitalized and enter state  $H$  with probability  $p^H$ . Otherwise, they go into state  $Q$ . This state  $Q$  is meant to hold those who self-quarantine after self-diagnosis or a positive test. However, the state also holds asymptomatic individuals and those who refuse to self-quarantine. Our model doesn't allow them to still infect others.

There are two reasons for this. The first reason is that it's hard to estimate the number of people in these categories.

The second reason is that asymptomatic infections are rare. A recent meta-analysis by Byambasuren et. al [2] estimated that 17% of disease carriers are asymptomatic, and that these carriers are 42% less likely to transmit the virus than a symptomatic carrier.

### 4.3 Reproduction rate and Mutation

In Schwarzendahl et al., [3], the inherent reproduction rate of the virus increases continuously, either linearly over time or in proportion to the current number of people infected. Instead, we propose that the reproduction rate  $\alpha_a(t)$  increases by  $\lambda$  after each variant appears. We assume that a variant appears once the global number of person-days infected reaches  $N$ . Thus, we are assuming that each person in state  $I$  contributes the same mutation risk (per day) and are assuming  $N$  is deterministic.

For a given SEIR model,  $N$  occurs at some time  $t_N$ . We assume that the mutation occurs in the area with the largest number of infected person-days, which we call area 1. In this area the variant becomes dominant after a delay of  $T$  days and the mixture of the two variants changes over time according to

$$\alpha_1(t) = a_0 + \frac{a_1 - a_0}{1 + e^{-k(t - (t_N + T_D))}} \quad (1)$$

Here  $k = \ln\left(\frac{1-p}{p}\right)/T$ , so that  $a_1(t_N) = a_0 + (a_1 - a_0)p$ , i.e.  $p$  is the proportion of the new variant at time  $t_N$ .

To visualize, see this temporary graphic:

Multiple variants should be able to occur. To be consistent with the linear (increasing) model in [3], more

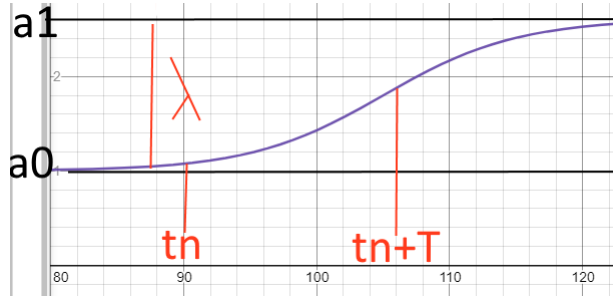


Figure 2: The infection rate  $\alpha(t)$  going through a mutation with parameters  $t_N = 90, T_D = 15, k = 0.2, \lambda = 1.6$

variants would occur, separated by  $N$  people-days of infection. These variants spread to other areas with a lag  $L$ , so that

$$\alpha_a(t) = \alpha_1(t + L) \text{ for } a \neq 1. \quad (2)$$

#### 4.4 Governmental and societal response

We use a unitless rate  $\gamma_a(t)$  as multiplier due to governmental and societal response, which varies with cases and then approaches a new normal.

Governmental and societal response is modeled as something like:

$$\gamma(t)_a = -\arctan\left(\frac{t}{k} - 1\right) + 1.5 \arctan\left(\frac{t}{100} - 1\right) + 2 \quad (3)$$

where  $k$  is the strength of the governmental/societal response, and 1.5 and 100 control the return to “new normal.”

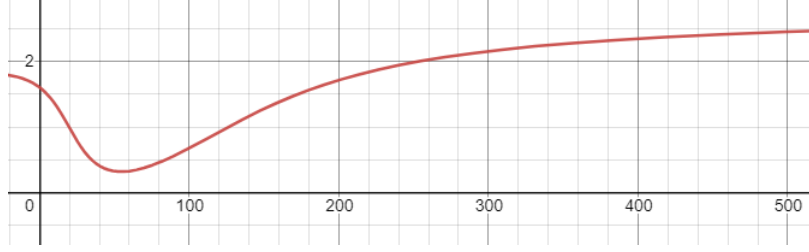


Figure 3:  $\gamma(t)$  with  $k = 20$

## 4.5 Mortality rate

The Johns Hopkins University Center for Systems Science and Engineering has [kept track of](#) COVID-19 mortality rates. Since the start of the pandemic, the mortality rate has decreased over time in all areas. After the first wave of the virus, the mortality rate in each area plateaued. It did not change much when the delta variant took over. It declined recently due to a younger average age and due to effective drugs that have been approved for treatment. We model mortality rate in area  $a$  by a constant  $m_a$ .

## 4.6 Vital Dynamics

We omit vital dynamics (birth and death) as a simplifying assumption. Instead, the sum of all people over all the SEIR states is held constant, even as the death state accumulates. See the appendix for how Vital Dynamics could be implemented.

## 4.7 SEIR Model Equations

$$\frac{dS_a}{dt} = -\beta V_a(t) - \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) \quad (4)$$

$$\frac{dE_a}{dt} = \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) - r^I E_a(t) \quad (5)$$

$$\frac{dI_a}{dt} = r^I E_a(t) - r^d I_a(t) \quad (6)$$

$$\frac{dQ_a}{dt} = r^d (1 - p^H) I_a(t) - r^R Q_a(t) \quad (7)$$

$$\frac{dH_a}{dt} = r^d p^H I_a(t) - r^R H_a(t) \quad (8)$$

$$\frac{dD_a}{dt} = r^R m_a [Q_a(t) + H_a(t)] \quad (9)$$

$$\frac{dR_a}{dt} = r^R (1 - m_a) [Q_a(t) + H_a(t)] \quad (10)$$

$$\frac{dM_a}{dt} = \beta V_a(t) \quad (11)$$

## 4.8 Future Work

The model could be adjusted to reflect the spread of the virus by asymptomatic individuals.

## 5 Vaccine optimization model

We formulate an optimization model to find the vaccine allocation to minimize deaths over a time horizon.

### Parameters:

$B(t)$  total vaccine doses available in time  $t$

$\rho_a$  proportion willing to be vaccinated in area  $a$

$\delta_a$  maximum that the fraction vaccinated per day in an area  $a$  can change per day

$f_a^{max}$  maximum fraction that can be vaccinated per day, per area  $a$ .

$T$  Time Horizon

For the optimization problem, we use a discrete-time version of the SEIR dynamics with a time step of one day. The difference equations are, e.g.,

$$M_a(t+1) = M_a(t) + \beta V_a(t) \quad (12)$$

The optimization model is

$$\begin{aligned}
& \min \sum_{a=1}^{|A|} D_a(T) && \text{(Deaths in time horizon)} \\
& \text{s.t. } \sum_{a=1}^{|A|} V_a(t) \leq B_t && t = 0, \dots, T-1 \quad \text{(Vaccination budget)} \\
& M_a(t) \leq \rho_a S_a(0) && t = 0, \dots, T-1, \quad a \in A \quad \text{(Vaccine willingness)} \\
& V_a(t) \leq f_a^{\max} S_a(0) && t = 0, \dots, T-1, \quad a \in A \quad \text{(Vaccine capacity)} \\
& |(V_a(t+1) - V_a(t))| \leq \delta_a S_a(0) && t = 0, \dots, T-1, \quad a \in A \quad \text{(Vaccine agility)} \\
& S_a(t+1) \geq S_a(t) - \beta V_a(t) - \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& E_a(t+1) \geq E_a(t) + \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) - r^I E_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& I_a(t+1) \geq I_a(t) + r^I E_a(t) - r^d I_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& Q_a(t+1) \geq Q_a(t) + r^d (1 - p^d p^H) I_a(t) - r^R Q_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& H_a(t+1) \geq H_a(t) + r^d p^H p^d I_a(t) - r^R H_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& D_a(t+1) = D_a(t) + r^R m_a [Q_a(t) + H_a(t)] && t = 0, \dots, T-1, \quad a \in A \\
& R_a(t+1) = R_a(t) + r^R (1 - m_a) [Q_a(t) + H_a(t)] && t = 0, \dots, T-1, \quad a \in A \\
& M_a(t+1) = M_a(t) + \beta V_a(t) && t = 0, \dots, T-1, \quad a \in A \\
& S_a(t), E_a(t), I_a(t), Q_a(t), H_a(t), V_a(t) \geq 0 && t = 0, \dots, T-1, \quad a \in A \quad \text{(Nonnegativity)}
\end{aligned}$$

Although the constraints are nonlinear, an iterative process can be used to solve this optimization problem as in (Bertsimas, et. al, 2021) [1]. In each iteration, given the current solution  $V$ , solve the difference equations to get  $\hat{I}_a(t)$ . In the (nonlinear) constraints for  $S$  and  $E$ , replace the variables  $I$  with the constants  $\hat{I}$ . Omit the constraints for  $I$  and add the regularization constraints

$$|I_a(t) - \hat{I}_a(t)| \leq \epsilon \quad t = 0, \dots, T-1, \quad a \in A \quad (13)$$

Here,  $\epsilon > 0$  is an error tolerance. This linear program can be solved to find the best solution  $V^*$  that gives infection dynamics  $I$  that are close to  $\hat{I}$ , i.e., close to those of the current solution. The current solution  $V$  is updated and the process repeated until the objective function value converges. Other stopping criteria could also be used, as in [1].

## 6 Uses of the model

This model is meant to describe COVID-19 on a global scale. One application would be to imagine four areas, as shown in table 1.

		Vaccinations	
		High	Low
Travel	High	Area 1	Area 2
	Low	Area 3	Area 4

Table 1: The four areas, defined by vaccination and travel rates.

Here are some of the countries in each area. It’s a work in progress, as the distinction between the areas is blurred. Data and code is documented in the [International Vaccine Allocation gitHhb](#).

area	Country	$\frac{\text{total vaccinations}}{\text{population}}\%$	$\frac{\text{total passengers}}{\text{population}}$	population
1	Canada	80.79	2.45	38,067,913
1	Ireland	77.5	34.14	4,982,904
1	Japan	79.22	1.03	126,050,796
1	United Arab Emirates	98.10	9.2	9,991,083
1	United Kingdom	74.87	2.08	682,071,14
1	United States	70.87	2.78	332,915,074
2	Bahamas	38.2	4.67	396,914
2	India	57.39	0.12	1,393,409,033
2	Indonesia	51.54	0.33	276,361,788
2	Mexico	59.39	0.53	130,262,220
2	Russia	46.60	0.79	145,912,022
2	South Africa	29.41	0.42	60,041,996
3	Argentina	81.07	0.42	45,605,823
3	Cambodia	83.55	0.08 8	16,946,446
3	Colombia	74.41	0.72	51,265,841
3	Cuba	90.02	0.040	11,317,498
3	Ecuador	77.12	0.26	17,888,474
3	Iran	67.49	0.25	85,028,760
4	Ethiopia	7.06	0.10	117,876,226
4	Ghana	8.36	0.019	31,732,128
4	Kenya	8.9	0.117	54,985,702
4	Mozambique	20.4	0.018	32,163,045
4	Nigeria	3.09	0.031	211,400,704
4	Uganda	8.15	0.0004	47,123,533

Table 2: Vaccination rates and flights per person for each country in each area

In this case, it would be useful to include the vaccine partial effectiveness assumption from 7.3. That way, vaccination prevents more people from contributing to mutation.

The interactions between these areas could shed insight on the effectiveness of vaccines to reduce mutations. That is, how beneficial is a vaccine when it is used on area 2? tic or movement of asymptomatic



## References

- [1] D. Bertsimas, J. Ivanhoe, A. Jacquillat, M. Li, A. Previero, O. S. Lami, and H. T. Bouardi, “Optimizing vaccine allocation to combat the covid-19 pandemic.,” *medRxiv*, 2020. [Available at <https://www.medrxiv.org/content/10.1101/2020.11.17.20233213v1>].
- [2] O. Byambasuren, M. Cardona, K. Bell, J. Clark, M.-L. McLaws, and P. Glasziou, “Estimating the extent of true asymptomatic covid-19 and its potential for community transmission: systematic review and meta-analysis,” 2020. Available at SSRN 3586675.
- [3] F. J. Schwarzendahl, J. Grauer, B. Liebchen, and H. Löwen, “Mutation induced infection waves in diseases like covid-19,” *medRxiv*, 2021.

## 7 Appendix: Model additions

### 7.1 Inter-area Travel

This addition accounts for the spread of the virus from travelling between areas. Infected travelers expose those in their area of destination.

**Parameters:**

$T_{a,l}$  Number of passengers on flights from area  $a$  to area  $l$  per year.

The median infectious person stays undetected for  $\frac{\ln(2)}{r^d}$  days, so that's their duration of infectiousness. The number of people from area  $a$  who are in area  $l$  who are also infectious is  $\frac{T_{a,l}}{365} \frac{I_a(t)}{N_a} \frac{\ln(2)}{r^d}$ . It's the number of travellers from  $a$  to  $l$  per day, times the proportion of those who are infectious, times their duration of infectiousness.

With this addition, the SEIR equations (4) and (5) become:

$$\frac{dS_a}{dt} = -\beta V_a(t) - \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} \left( I_a(t) + \sum_{l=1}^{|A|} \frac{T_{l,a}}{365} \frac{I_l}{N_l} \frac{\ln(2)}{r^d} - \sum_{a=1}^{|A|} \frac{T_{a,l}}{365} \frac{I_a}{N_a} \frac{\ln(2)}{r^d} \right) \quad (14)$$

$$\frac{dE_a}{dt} = \alpha_a(t) \gamma_a(t) \frac{S_a(t)}{N_a} \left( I_a(t) + \sum_{l=1}^{|A|} \frac{T_{l,a}}{365} \frac{I_l}{N_l} \frac{\ln(2)}{r^d} - \sum_{a=1}^{|A|} \frac{T_{a,l}}{365} \frac{I_a}{N_a} \frac{\ln(2)}{r^d} \right) - r^I E_a(t) \quad (15)$$

The term  $I_a(t) + \sum_{l=1}^{|A|} \frac{T_{l,a}}{365} \frac{I_l}{N_l} \frac{\ln(2)}{r^d} - \sum_{a=1}^{|A|} \frac{T_{a,l}}{365} \frac{I_a}{N_a} \frac{\ln(2)}{r^d}$  represents the people infected, plus the out of country infectious people in  $a$ , less the infectious people from  $a$  who are in other areas.

The SEIR equations (4)-(11) for the other states are unchanged.

### 7.2 Vital Dynamics

Vital dynamics (births and non-pandemic related deaths) are significant when simulating a long time horizon.

**Parameters:**

$\mu$  Death rate (proportion of population per day)

We make the simplifying assumption that the number of births equals the number of natural deaths. We add the births,  $\mu N_a$ , into the susceptible state. Deaths leave all other states except  $D$ , e.g.  $-\mu S_a(t)$  for state  $S_a$ .

A simplified diagram shows the addition of vital dynamics. In this simple version,  $N_a = S_a + I_a + R_a$ .

The full model, with vital dynamics, would be:

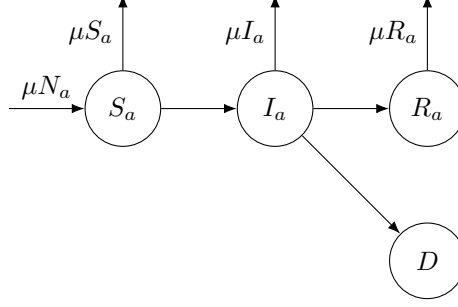


Figure 4: The dynamics of births and deaths on a simple SIR model

$$\frac{dS_a}{dt} = -\beta V_a(t) - \alpha_a(t)\gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) + \mu N_a - \mu S_a \quad (16)$$

$$\frac{dE_a}{dt} = \alpha_a(t)\gamma_a(t) \frac{S_a(t)}{N_a} I_a(t) - r^I E_a(t) - \mu E_a \quad (17)$$

$$\frac{dI_a}{dt} = r^I E_a(t) - r^d I_a(t) - \mu I_a \quad (18)$$

$$\frac{dQ_a}{dt} = r^d(1 - p^d p^H) I_a(t) - r^R Q_a(t) - \mu Q_a \quad (19)$$

$$\frac{dH_a}{dt} = r^d p^H p^d I_a(t) - r^R H_a(t) - \mu H_a \quad (20)$$

$$\frac{dD_a}{dt} = r^R m_a [Q_a(t) + H_a(t)] \quad (21)$$

$$\frac{dR_a}{dt} = r^R(1 - m_a) [Q_a(t) + H_a(t)] - \mu R_a \quad (22)$$

$$\frac{dM_a}{dt} = \beta V_a(t) - \mu M_a \quad (23)$$

### 7.3 Vaccine Partial effectiveness

It is more realistic for the vaccine to give people partial immunity rather than giving a proportion of people full immunity. Here, everyone who is vaccinated at time  $t$  in area  $a$ ,  $V_a(t)$ , is brought into a new set of states.

#### State Variables:

$S_a^V$  Susceptible and vaccinated

$E_a^V$  Exposed and vaccinated

$I_a^V$  Infectious and vaccinated

These states are added to figure 1.

#### Parameters:

$p^e$  Vaccine effectiveness at reducing transmission

$p^r$  Vaccine protection against receiving the virus

As you can see in Figure 5, those who are vaccinated can get sick but they can't die. This assumption is

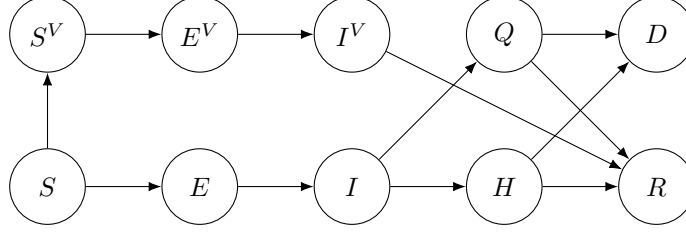


Figure 5: The states for an area  $a$  with a vaccination track

due to the observed reduced mortality of breakthrough cases. The population  $I^V$  is less infectious than the state  $I$  due to the vaccine's effectiveness against transmission: an infectious person who has been vaccinated is only  $1 - p^e$  as infectious as an unvaccinated individual. Those who are vaccinated are less likely to receive the virus. We represent this with the proportion  $p^r$ .

#### 7.4 SEIR Model with vaccination

$$\frac{dS_a}{dt} = -V_a(t) - \alpha_a(t)\gamma_a(t)\frac{S_a(t)}{N_a}[I_a(t) + (1 - p^e)I_a^V(t)] \quad (24)$$

$$\frac{dE_a}{dt} = \alpha_a(t)\gamma_a(t)\frac{S_a(t)}{N_a}[I_a(t) + (1 - p^e)I_a^V(t)] - r^I E_a(t) \quad (25)$$

$$\frac{dI_a}{dt} = r^I E_a(t) - r^d I_a(t) \quad (26)$$

$$\frac{dS_a^V}{dt} = V_a(t) - p^r \alpha_a(t)\gamma_a(t)\frac{S_a^V(t)}{N_a}[I_a(t) + (1 - p^e)I_a^V(t)] \quad (27)$$

$$\frac{dE_a^V}{dt} = p^r \alpha_a(t)\gamma_a(t)\frac{S_a^V(t)}{N_a}[I_a(t) + (1 - p^e)I_a^V(t)] - r^I E_a^V(t) \quad (28)$$

$$\frac{dI_a^V}{dt} = r^I E_a^V(t) - r^d I_a^V(t) \quad (29)$$

$$\frac{dQ_a}{dt} = r^d(1 - p^d p^H)I_a(t) - r^R Q_a(t) \quad (30)$$

$$\frac{dH_a}{dt} = r^d p^H p^d I_a(t) - r^R H_a(t) \quad (31)$$

$$\frac{dD_a}{dt} = r^R m_a [Q_a(t) + H_a(t)] \quad (32)$$

$$\frac{dR_a}{dt} = r^R(1 - m_a)[Q_a(t) + H_a(t)] \quad (33)$$

We assume that breakthrough cases do not contribute to mutation because the virus replicates less. Hence,  $t_N$  is computed from  $I$ , not  $I^V$ .