UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS
Y DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

Probabilistic epidemiological model for heterogeneous populations

TESINA
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
JOSÉ GIRAL BARAJAS

DOCTOR SERGIO IVÁN LÓPEZ ORTEGA
FACULTAD DE CIENCIAS
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

CIUDAD UNIVERSITARIA, CD. MX., JUNIO 2024

Introducción

En esta tesina, presentamos una extensión de un nuevo modelo epidemiológico estocástico basado en procesos de muestreo para considerar poblaciones heterogéneas. En la primera sección, introducimos el problema y el trabajo existente. En la segunda sección, definimos la extensión del modelo para poblaciones heterogéneas y presentamos diferentes formas de abordar el problema. En la tercera sección, definimos el número reproductivo básico del modelo y evaluamos algunas de sus propiedades. Finalmente, en la última sección, presentamos una aplicación del modelo al caso de COVID-19, considerando las variantes Ómicron y Delta y el impacto de la transmisión asintomática.

Introduction

In this short dissertation, we introduce an extension of a new stochastic epidemiological model based on sampling processes to consider heterogeneous populations. In the first section, we introduce the problem and the existing work. In the second section, we define the extension of the model for heterogeneous populations and present different ways of tackling the problem. In the third section, we define the basic reproductive number of the model and evaluate some of its properties. Finally, in the last section, we present an application of the model to the case of COVID-19, considering Omicron and Delta variants and the impact of asymptomatic transmission.

1 Preliminaries

Throughout history, outbreaks of various infectious diseases have significantly impacted humanity. From the devastating black plague in the 14th century to the unprecedented challenges posed by the COVID-19 pandemic in the 21st century, these outbreaks have exacted an enormous toll of human lives and in many socio-economic indicators (Britton et al., 2019). Therefore, a primary task for public health is to understand how infectious diseases spread throughout populations to predict, control, and ultimately eradicate possible outbreaks. Scientific experimentation in this context is complex to design, expensive to implement, and often falls into unethical territory. That is why mathematical models are one of the best tools for this task.

Epidemiological models allow us to optimise resources in the study of epidemic outbreaks and evaluate the effectiveness of possible control measures by understanding and predicting epidemiology dynamics at the population level. The approach for modelling can vary depending on several factors, depending on the objectives of the study. For example, there are models focused on predicting hospital bed occupancy, evaluating the effectiveness of possible mitigation measures, or understanding the influence of asymptomatic transmission.

Lying at the foundation of modern epidemiological modelling is the family of SIR-type models (Allen, 2010; Brauer and Castillo-Chavez, 2012) presented in a seminal work by Kermack and McKendrick (Kermack and McKendrick, 1927). These models divide the population into three compartments labelled S, I, and R. We let the non-negative real numbers $S(t)$, $I(t)$, and $R(t)$ respectively denote the proportion of the population that is susceptible to disease, infected, and removed from the epidemic process at time t. Removed individuals can be thought of as isolated from the rest of the population or as totally and permanently immune to infection.

The simplest form of this model has several assumptions worth remarking. First, it considers a closed population with no migration, births, or deaths due to other factor. Therefore, population size is constant, i.e. $S(t) + I(t) + R(t) = 1$ at all time t. Second, the model assumes homogeneous mixing: everyone can have contact with anyone else with the same probability, and individuals are indistinguishable from one another; this translates to a mean-field assumption that gives rise to the nonlinear dynamics of the model. Finally, the model is defined in continuous time ($t \in \mathbb{R}_+$) and with population proportions ($S(t), I(t), R(t) \in [0, 1]$ for all $t \geq 0$). This model is determined by the following set of three ordinary differential equations

$$\begin{aligned} \frac{dS}{dt} &= -\beta SI, \\ \frac{dI}{dt} &= \beta SI - \gamma I, \\ \frac{dR}{dt} &= \gamma I, \end{aligned} \tag{1.1}$$

where β is the transmission rate, also thought as the transmission probability given an infectious contact, and γ is the removal rate. Observe that, as the population is constant, $R(t) = 1 - S(t) - I(t)$ and the system (1.1) can be reduced to a system of two equations.

The SIR model (1.1) has a disease-free equilibrium when $S = 1$ and $I = R = 0$. We can also observe that $\frac{dS}{dt} \leq 0$ for every t as $S(t)$, $I(t)$, and β are always non-negative.

Furthermore, $\frac{dI}{dt} > 0$ if and only if $S(t) > \gamma/\beta$. Therefore, if $S(0) < \gamma/\beta$, the proportion of infections $I(t)$ decreases to zero and there is no epidemic outburst. On the other hand, if $S(0) > \gamma/\beta$, the proportion of infections $I(t)$ increases until it reaches a maximum, when $S(t) = \gamma/\beta$, and then decreases to zero, constituting an epidemic outburst (Brauer and Castillo-Chavez, 2012). With this insight, one can define the threshold quantity $\mathcal{R}_0 := \beta/\gamma$, called the *basic reproductive number*, that determines whether there will be an epidemic outburst. In that sense, if $\mathcal{R}_0 > 1$, there will be an outburst, while if $\mathcal{R}_0 < 1$, the pathogen infection will vanish from the host population.

Recently, we proposed a simple probabilistic model for epidemiology dynamics, hereby called sED model (Giral-Barajas et al., 2023). This model recovers different types of deterministic SIR models, builds an extension that takes into account a more realistic description of the disease history of individuals, and obtains case-fatality ratios that are closer to those observed in reality (Giral-Barajas et al., 2023). The model is based on the idea that transitions between epidemiological stages are similar to sampling processes and are dependent on the elapsed time and routing probabilities towards recovery or deterioration as defined by pathological or clinical criteria. The basic model assumes that the number $X_I^h(t)$ of newly infected people is a *binomial* random variable sampled from $S(t)$ individuals, extracted from a total population with size $T(t)$, with probability

$$p^h(t, I) = P_I(h)\beta\frac{I(t)}{T(t)}, \tag{1.2}$$

where β is the probability of infection given an infectious contact, $I(t)/T(t)$ represents the proportion of infected individuals at time t, and $P_I(h)$ represents the probability that the infection occurs within a time window of length h.

Following infection, the model assumes that some individuals die after becoming infected, and the rest recover. $Y_R^h(t)$ and $Y_D^h(t)$ represent the numbers of people that recover or die respectively between t and $t + h$. The probability that a person is removed from the infected group is assumed to depend only on the physiological state of the individual facing the infection. To model this progression, we assumed that infected individuals can either recover within an average time μ_R or die from disease after an average time μ_D. Consequently, the probabilities that a person is removed from the infected group within a time interval of length h because of recovery or death can be estimated as

$$p_s^h \approx h\,\mu_s^{-1}, \quad s \in \{R, D\}. \tag{1.3}$$

Therefore, the probability of remaining infected within h time units is $(1 - p_R^h + p_D^h)$.

The number of newly recovered and deceased people $Y_R^h(t)$ and $Y_D^h(t)$ can be obtained as outcomes from a *Multinomial* sample

$$(Y_R^h(t), Y_D^h(t), I(t) - Y_R^h(t) - Y_D^h(t)) \sim \text{Mult}(I(t), p_R^h, p_D^h, 1 - p_R^h - p_D^h). \tag{1.4}$$

The dynamics of the epidemiological components are described by the following stochastic dynamical system, defining the basic sED model,

$$\begin{aligned} S(t+h) &= S(t) - X_I^h(t), \\ I(t+h) &= I(t) + X_I^h(t) - (Y_R^h(t) + Y_D^h(t)), \\ R(t+h) &= R(t) + Y_R^h(t), \\ D(t+h) &= D(t) + Y_D^h(t). \end{aligned} \tag{1.5}$$

This basic model allows modelling infection spread in small populations with integer state variables while maintaining assumptions of a closed population and homogeneous mixing. On top of that, it falls short on describing accurate case-fatality ratios (CFRs) for an epidemic coming from infection with shorter hospitalisation times with fatal outcomes in comparison to recovery times, as was reported during the beginning of the COVID-19 pandemic (Zhou et al., 2020). To address this issue, we proposed an adaptation of the basic sED, with the main idea being to take into account a more realistic description of the disease history of individuals. To achieve this, the time course of infection was divided into different stages that progress towards more severe illness until possibly reaching death in the absence of recovery (Fig. 2.1 B). In other words, at each infectious stage, the individual either recovers or advances to a more severe state. At the last infectious stage, recovery or death are the only possibilities. We omit the equations as they resemble the ones derived in the following section and can be found in the original study (Giral-Barajas et al., 2023).

The dynamics from this extension of the basic sED model allow the expected time before recovery to be larger than the expected time before death and still obtain CFRs similar to those observed in the data of the COVID-19 pandemic. Nevertheless, as the pandemic gained ground, significant variability in its progression could be seen between countries (Hradsky and Komarek, 2021). Even in small groups of people sharing some social, geographical, and biological characteristics, the heterogeneity produced by a non-shared characteristic resulted in a wide range of histories of infection and outcomes (Pereira et al., 2021). Such variability in the development of disease exhibits the urgency in understanding the epidemiology dynamics while considering the diversity of social and biological factors that characterise individuals.

2 Epidemiological model for heterogeneous populations

The following section presents an original extension of the basic and multistage sED models (Giral-Barajas et al., 2023). This extension aims to study the epidemiology dynamics of an infectious disease in a heterogeneous population. We start by diving the total population in N sub-populations characterised by m traits. Let us consider the vector $v = (v_1, \ldots, v_m) \in \mathbb{R}^m$, where each entry describes one biological or socially identifiable characteristic. For example, v_i could represent an individual's age in years, taking values in a finite, bounded set of possible ages. An alternative would be to divide the total population into age groups, and v_i would represent the age group. A second example could be that v_i represents a previously diagnosed medical condition that constitutes a comorbidity in the case of an infection. For instance, we could take $v_i = 1$ if the person has been previously diagnosed with diabetes or $v_i = 0$ if not. A final example, now of a social characteristic, would be to consider $v_i = k$ for $k \in \{1, 2, \ldots, l\}$, for $l \in \mathbb{N}$, where k represents the proximity to medical facilities.

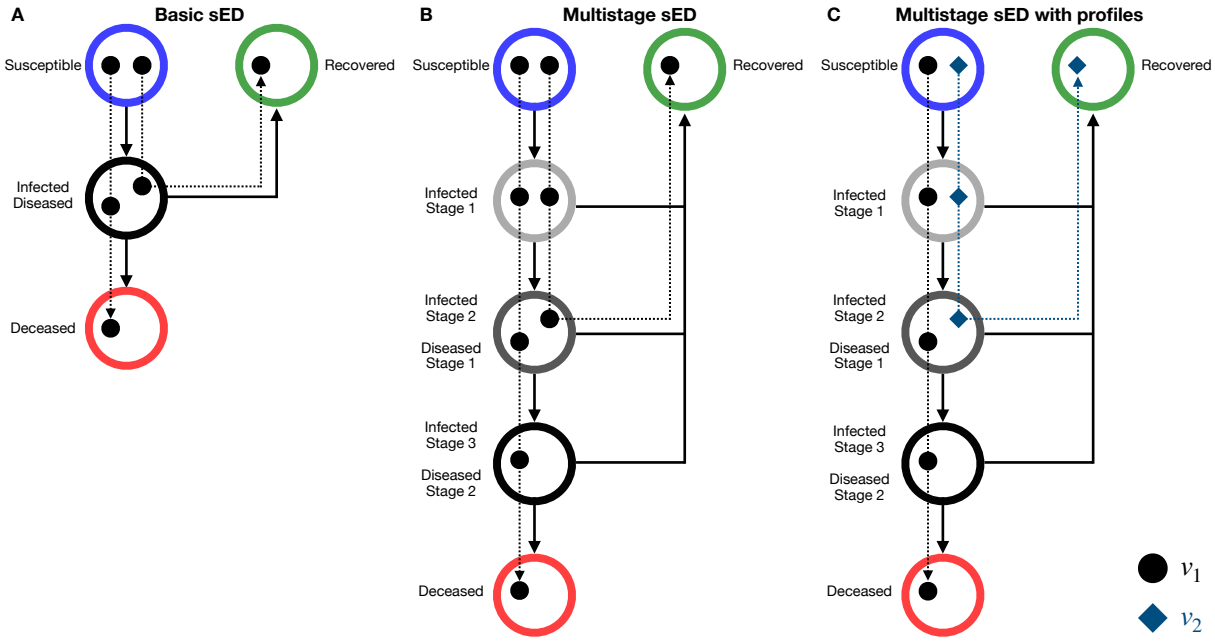

Figure 2.1: Different dynamical scenarios contemplated by the sED modelling scheme. Basic sED and multistage sED are outlined in A and B (Giral-Barajas et al., 2023). The new multistage sED with two profiles (v_1 and v_2) is outlined in C.

The characteristics represented in v can be categorical, ordinal or interval data types. For the sake of simplicity, as each of these data types is a subset of $\mathbb{R}$, we define v as an element of $\mathbb{R}^m$. From now on, we will refer to the vector $v = (v_1, \ldots, v_m)$, which lists all the characteristics of an individual, as a *profile*, and denote $\mathcal{P}$ the *set of all possible profiles*. It follows that $\mathcal{P} \subseteq \mathbb{R}^m$ and $\mathcal{P}$ is finite with cardinality N. Therefore, at each time t, the total population can be divided into N sub-populations of individuals with the same profile. That is

$$T(t) = \sum_{v \in \mathcal{P}} \pi^v(t) T(t), \tag{2.1}$$

where $\pi^v(t)$ represents the proportion of individuals with profile v at time t.

We can construct an extension of the multistage sED model with these basic definitions. The central assumption is that individuals with similar profiles will have similar trajectories relative to infection. Examples of communicable diseases of current importance where individuals have similar clinical progressions depending on some of their characteristics include dengue (Sirisena et al., 2021), COVID-19 (Zhao et al., 2019), tuberculosis (Trauer et al., 2019), among many well-documented examples. In this way, if two persons share a profile, they will have the same parameters associated with the evolution of a multistage sED model; this implies that within each subpopulation, there is homogeneity. As individual infections progress according to a multistage sED, the time course of infection is divided into different stages that progress towards more severe illness until possibly reaching death in the absence of recovery. Therefore, for each profile $v \in \mathcal{P}$, we represent the sizes of the infected populations at time t by k variables $I_0^v(t), \ldots, I_{k-1}^v(t)$ with increasing indexes according to the severity of disease (Fig. 2.1 C). We begin by describing how to model the contagion of susceptible individuals.

2.1 Transition from susceptible to infected

In general, we assume that new infections occur randomly and independently among susceptible individuals. The probability of infection may depend on factors such as the duration of exposure to infection h and the amount of available inoculum. With this paradigm, we now have two possible ways of modelling the transition from susceptible to infected, which will depend on the profiles of the individuals involved.

2.1.1 Static profiles

In a first model, profiles remain fixed over time and are determined a priori within the population. We portray characteristics such as age, presence of comorbidities, and socioeconomic status, among others. An individual does not acquire the profile from another individual as a result of infection. Thus, the population partition in equation (2.1) is preserved in all epidemiological stages, including the susceptible population. Consequently, for each profile $v \in \mathcal{P}$, the non-negative integer $S^v(t)$ represents the susceptible population with profile v at time t (in days).

As infections are independent among individuals, the number of newly infected people with profile v within a time window of size h at time t, denoted by $X_0^{v,h}(t)$, can be thought as a binomial random variable sampled over the $S^v(t)$ susceptible individuals of said profile. One possibility is to define the average probability of infection as

$$p_h^v(I_0, \ldots, I_{k-1}, t) = h\beta^v \sum_{v' \in \mathcal{P}} \sum_{i=0}^{k-1} \kappa_i^{v'}(t) \frac{I_i^{v'}(t)}{T(t)}, \tag{2.2}$$

where β^v is the probability that a susceptible person with profile v gets infected after having an infectious contact and $\kappa_i^{v'}(t)$ represents the proportion of infected individuals at i-th stage of infection with profile v' that are exposed for an infectious contact with susceptible individuals. Let us remark that the probability that a susceptible person gets infected after having an infectious contact depends on their profile through the parameter β^v. In such a way, we can interpret β^v as a parameter of how susceptible an individual with profile v is to the infectious pathogen. Therefore, for each profile $v \in \mathcal{P}$ the sampling underlying the evolution of the newly infected is given by

$$X_0^{v,h}(t) \sim Bin(\lfloor \varepsilon^v(t) S^v(t) \rfloor, p_h^v(I_0, \ldots, I_{k-1}, t)), \tag{2.3}$$

where $\varepsilon^v(t)$ represents the proportion of exposed susceptible individuals with profile v. On that account, the number of susceptible individuals with profile $v \in \mathcal{P}$ at time $t + h$ and in the whole population can be computed as

$$S^v(t+h) = S^v(t) - X_0^{v,h}(t) \quad \text{y} \quad S(t+h) = \sum_{v \in \mathcal{P}} S^v(t+h), \tag{2.4}$$

respectively. As an alternative model, we define another type of profile that modifies the sampling processes determining the transition from susceptible to infected.

2.1.2 Dynamic profiles

In this model, an individual's profile may change over time and is initially determined during infection. We can use this type of profile to model characteristics such as the variant of a pathogen with which an individual was infected, or the number of times they have been infected. In contrast with the previous model, an individual acquires a new profile during contagion. In this case, the population partition in equation (2.1) is preserved for all epidemiological stages except for the susceptible population. The non-negative integer $S(t)$ represents the size of the whole susceptible population at time t, and there is no heterogeneity in susceptibility as none of the susceptible individuals start with a defined profile.

Analogously to static profiles, the number of newly infected people with profile v within a time window of size h at time t, denoted by $X_0^{v,h}(t)$, can be thought of as a binomial random variable sampled over the susceptible individuals with average probability of infection

$$p_h^v(I_0, \ldots, I_{k-1}, t) = h \sum_{i=0}^{k-1} \beta_i^v \kappa_i^v(t) \frac{I_i^v(t)}{T(t)}, \tag{2.5}$$

where β_i^v is the probability that a susceptible person gets infected after having contact with an infected person with profile v at the i-th stage of infection and $\kappa_i^v(t)$ represents the proportion of infected individuals at i-th stage of infection with profile v that are exposed for an infectious contact with susceptible individuals. We want to remark that, in this case, the probability that a susceptible person gets infected after having an infectious contact depends exclusively on the profile of the contagious person with whom they had contact. In such a way, we can interpret β_i^v as a parameter of how infectious the pathogen is coming from an individual with profile v at infectious stage i. As there is only one group of susceptible individuals, the sampling underlying the dynamics must be carried out simultaneously. So we have that

$$\Big(X_0^{v_1,h}(t), \ldots, X_0^{v_N,h}(t), S(t) - \sum_{v \in \mathcal{P}} X_0^{v,h}(t)\Big) \sim Mult\Big(\lfloor \varepsilon(t) S(t) \rfloor, p_h^{v_1}(I,t), \ldots, p_h^{v_N}(I,t), 1 - \sum_{v \in \mathcal{P}} p_h^v(I,t)\Big), \tag{2.6}$$

where $\varepsilon(t)$ represents the proportion of susceptible individuals exposed to infection. In this way, for each profile $v \in \mathcal{P}$, the sampling underlying the evolution of the newly infected is given by

$$X_0^{v,h}(t) \sim Bin(\lfloor \varepsilon(t) S(t) \rfloor, p_h^v(I_0, \ldots, I_{k-1}, t)), \tag{2.7}$$

and the number of susceptible individuals at time $t+h$ in the whole population can be computed as

$$S(t+h) = S(t) - \sum_{v\in\mathcal{P}} X_0^{v,h}(t). \tag{2.8}$$

A limitation of these profile definitions lies in the probability of infection given an infectious contact. On the one hand, with static profiles, this probability only depends on the profile of the susceptible and not on the severity of those infected. On the other hand, with dynamic profiles, we have a homogeneous susceptible population that is equally vulnerable to infection. We now describe the progression of an individual through disease classes once they have been infected.

2.2 Progression of infection through multiple stages

As mentioned earlier, we will represent with k variables $I_0^v(t), \ldots, I_{k-1}^v(t)$ the sizes of the infected populations at time t for each profile $v \in \mathcal{P}$. Indexes reflect severity of infection and disease. At each time step, each infected individual may move forward to the next infectious stage, recover from the disease, or stay at the same stage. In analogy with the multistage sED model, we will assume that the probabilities that a person with profile v in the i-th infectious stage advances to the next stage or recovers, within a small time interval of length h, are estimated by

$$p_I^{(v,i)} \approx \frac{h}{\mu_I^{(v,i)}} \quad \text{y} \quad p_R^{(v,i)} \approx \frac{h}{\mu_R^{(v,i)}}, \tag{2.9}$$

for $i \in \{0, \ldots, k-1\}$. Here, for a patient with profile v, $\mu_I^{(v,i)}$ is the mean time they spend in the i-th infectious stage before moving forward and $\mu_R^{(v,i)}$ is the mean time they spend in the i-th infectious stage before recovering (Giral-Barajas et al., 2023). The probability that an individual with profile v remains in the i-th infectious stage within h time units is $(1 - p_R^{(v,i)} - p_I^{(v,i)})$.

To simplify the notation, we identify the number of individuals who die due to disease at time t, denoted by $D(t)$, with an extra infectious class $I_k(t)$. For each profile $v \in \mathcal{P}$, disease progression dynamics are based on the underlying multinomial samplings, and

$$\begin{aligned}
I_j^v(t+h) &= I_j^v(t) + X_j^{v,h}(t) - \Big(X_{j+1}^{v,h}(t) + Y_j^{v,h}(t)\Big), && (2.10)\\
R_j^v(t+h) &= R_j(t) + Y_j^h(t), && (2.11)\\
D_j^v(t+h) &= D^v(t) + X_k^{v,h}(t), && (2.12)
\end{aligned}$$

for $j \in \{0, \ldots, k-1\}$. The total number of infected, recovered, and deceased is given by

$$I(t+h) = \sum_{v \in \mathcal{P}} \sum_{j=0}^{k-1} I_j^v(t+h), \tag{2.13}$$

$$R(t+h) = \sum_{v \in \mathcal{P}} \sum_{j=0}^{k-1} R_j^v(t+h), \tag{2.14}$$

$$D(t+h) = \sum_{v \in \mathcal{P}} D^v(t+h), \tag{2.15}$$

where $R^v(t)$ and $D^v(t)$ are the numbers of recovered and deceased individuals with profile v at time t, and $I(t), R(t)$ y $D(t)$ are the total number of infected, recovered, and deceased at time t. The sampling underlying disease progression is given by

$$(X_{j+1}^{v,h}(t), Y_j^{v,h}(t), I_j^v(t) - X_{j+1}^{v,h}(t) - Y_j^{v,h}(t)) \sim Mult\big(I_j^v(t), p_I^{(v,j)}, p_R^{(v,j)}, 1 - p_I^{(v,j)} - p_R^{(v,j)}\big), \tag{2.16}$$

for $j \in \{0, \ldots, k-1\}$. Observe that $X's$ represent the number of people that advance through infectious stages, while $Y's$ represent the number of people that recover from disease.

3 Basic reproductive number

As we saw in the introduction, the reproductive number for the classic SIR model determines whether there is an epidemic outburst or not, and arises from analysing when the instantaneous rate of change of infected people is positive, $\frac{dI}{dt} > 0$. This section will extend this threshold quantity for the multistage sED model with profiles. First, observe that, unlike the deterministic SIR, the multistage sED model with profiles is conformed by random variables that take integer values and progress in discrete time. Therefore, instead of analysing the infinitesimal increment of infected people, we will analyse the expected discrete increment given by $\mathbb{E}\left[\frac{\Delta_h I(t)}{h}\right]$ where $\Delta_h I(t) = I(t+h) - I(t)$. Taking in consideration equations (2.10) and (2.13), it follows that $\Delta_h I(t)$ is given by

$$\Delta_h I(t) = \sum_{v \in \mathcal{P}} \sum_{j=0}^{k-1} (X_j^{v,h}(t) - X_{j+1}^{v,h}(t) - Y_j^{v,h}(t)) = \sum_{v \in \mathcal{P}} \Big(X_0^{v,h}(t) - X_k^{v,h}(t) - \sum_{j=0}^{k-1} Y_j^{v,h}(t) \Big).$$

Therefore, we have that the average change of infected people is given by

$$\mathbb{E}\left[\Delta_h I(t)\right] = \sum_{v\in\mathcal{P}} \Big(\mathbb{E}\left[X_0^{v,h}(t)\right] - \mathbb{E}\left[X_k^{v,h}(t)\right] - \sum_{j=0}^{k-1} \mathbb{E}\left[Y_j^{v,h}(t)\right] \Big). \tag{3.1}$$

Observe that $\mathbb{E}\left[X_0^{v,h}(t)\right]$ depends on the type of profiles used in the model (dynamic or static) but $\mathbb{E}\left[X_k^{v,h}(t)\right]$ and $\mathbb{E}\left[Y_j^{v,h}(t)\right]$ do not, as $X_k^{v,h}(t)$ and $Y_j^{v,h}(t)$ are determined by equation (2.16), which does not take into account profile type. Therefore, we will develop the analysis without specifying the type of profiles being modelled until the end of the computations. In this way, considering equations (2.9) and (2.16), we have that the expected change of infected people is given by

$$\mathbb{E}\left[\Delta_h I(t)\right] = \sum_{v\in\mathcal{P}} \Big(\mathbb{E}\left[X_0^{v,h}(t)\right] - \frac{h\mathbb{E}\left[I_{k-1}^v(t)\right]}{\mu_I^{(v,k-1)}} - \sum_{j=0}^{k-1} \frac{h\mathbb{E}\left[I_j^v(t)\right]}{\mu_R^{(v,j)}} \Big). \tag{3.2}$$

Now, the expectations of this equation are not easily calculated directly, since $X_0^{v,h}(t)$ and $I_0^v(t), \ldots, I_{k-1}^v(t)$ are coupled stochastic processes that depend on these values at previous times. However, at the beginning of the epidemic, when evaluating these processes at $t = 0$, these quantities are known and determined as initial conditions. Furthermore, by evaluating at $t = 0$ we are studying epidemic outbreaks when the population is largely susceptible. For those reasons, we evaluate equation (3.2) at $t = 0$, assuming that the population is largely susceptible and that there are no individuals in advanced stages of infection, that is $I_j(0) = 0$ for every $j \neq 0$. Consequently

$$\mathbb{E}\left[\Delta_h I(0)\right] = \sum_{v\in\mathcal{P}} \Big(\mathbb{E}\left[X_0^{v,h}(0)\right] - \frac{hI_0^v(0)}{\mu_R^{(v,0)}} - \mathbb{1}_{\{k=1\}} \frac{hI_0^v(0)}{\mu_I^{(v,0)}} \Big), \tag{3.3}$$

where the indicator function appears in the last term since, in the case of a single stage of infection, infected individuals who die should be taken into account for the total change in infected people. Finally, dividing by h and using linearity of the expectation, we obtain that the expected rate change of infected people at the beginning of the epidemic is

$$\mathbb{E}\left[\frac{\Delta_h I(0)}{h}\right] = \sum_{v\in\mathcal{P}} \Big(\frac{1}{h}\mathbb{E}\left[X_0^{v,h}(0)\right] - \frac{I_0^v(0)}{\mu_R^{(v,0)}} - \mathbb{1}_{\{k=1\}} \frac{I_0^v(0)}{\mu_I^{(v,0)}} \Big). \tag{3.4}$$

Thus, for the expected rate of change in $I(0)$ to be positive, we obtain the following condition

$$\mathcal{R}_0 := \frac{1}{\sum_{v\in\mathcal{P}} I_0^v(0)\left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)} \cdot \frac{1}{h}\sum_{v\in\mathcal{P}} \mathbb{E}\left[X_0^{v,h}(0)\right] > 1. \tag{3.5}$$

From now on, we will refer to this quantity, $\mathcal{R}_0$, as the *general expression of the basic reproductive number for the multistage sED model with profiles* or simply *general* $\mathcal{R}_0$. As mentioned before, the value of $\mathbb{E}\left[X_0^{v,h}(0)\right]$ depends on the type of profile used in the model. Therefore, we will analyse the general expression of the basic reproductive number for models with each type of profile.

Static profiles First, in the case of the static profiles (subsection 2.1.1), it follows from equations (2.2) and (2.3) that for each profile $v \in \mathcal{P}$

$$\mathbb{E}\left[X_0^{v,h}(0)\right] = \lfloor \varepsilon^v(0)S^v(0)\rfloor h\beta^v \sum_{v'\in\mathcal{P}} \kappa_0^{v'}(0)\frac{I_0^{v'}(0)}{T(0)}.$$

Therefore, in this case, we have that the general expression of the basic reproductive number takes the following value

$$\mathcal{R}_0^s = \frac{\sum_{v\in\mathcal{P}} \lfloor \varepsilon^v(0)S^v(0)\rfloor \beta^v \sum_{v'\in\mathcal{P}} \kappa_0^{v'}(0)\frac{I_0^{v'}(0)}{T(0)}}{\sum_{v\in\mathcal{P}} I_0^v(0)\left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)}. \tag{3.6}$$

From now on, we will refer to this particular expression of the general $\mathcal{R}_0$, denoted by $\mathcal{R}_0^s$, as *basic reproductive number for the multistage sED model with static profiles* or simply $\mathcal{R}_0$ *for static profiles.*

Dynamic profiles Considering now dynamic profiles (subsection 2.1.2), it follows from equations (2.5) and (2.6) that for each profile $v \in \mathcal{P}$

$$\mathbb{E}\left[X_0^{v,h}(0)\right] = \lfloor \varepsilon(0)S(0)\rfloor h\beta_0^v\kappa_0^v(0)\frac{I_0^v(0)}{T(0)}.$$

Thus, in this case, we have that the general expression of the basic reproductive number boils down to

$$\mathcal{R}_0^d = \frac{\sum_{v\in\mathcal{P}} \lfloor \varepsilon(0)S(0)\rfloor \beta_0^v\kappa_0^v(0)\frac{I_0^v(0)}{T(0)}}{\sum_{v\in\mathcal{P}} I_0^v(0)\left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)}. \tag{3.7}$$

In line with previous cases, from now on, we will refer to this particular expression of the general $\mathcal{R}_0$, denoted by $\mathcal{R}_0^d$, as *basic reproductive number for the multistage sED model with dynamic profiles* or simply $\mathcal{R}_0$ *for dynamic profiles*.

As one can observe from equations (3.6) and (3.7), the basic reproductive number with static profiles and the basic reproductive number with dynamic profiles have similar expressions; the only difference lies in the numerator. Note that if $\beta^v = \beta$ for all $v \in \mathcal{P}$ in equation (3.6) and $\beta_0^v = \beta$ for all $v \in \mathcal{P}$ in equation (3.7) then we obtain the same expression for the basic reproductive number with dynamic and static profiles. This expression is given by

$$\mathcal{R}_0 = \frac{\beta \sum_{v\in\mathcal{P}} \lfloor \varepsilon(0)S(0) \rfloor \kappa_0^v(0) \frac{I_0^v(0)}{T(0)}}{\sum_{v\in\mathcal{P}} I_0^v(0) \left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)}.$$

Therefore, if we are modelling the dynamics of an infectious disease for which the probability of infection given an infectious contact is the same for all individuals in the population, we would have the same basic reproductive number regardless of the type of profile used. This is because in this case, profiles do not have relevance at the beginning of the epidemic.

3.1 Multistage sED and deterministic SIR-type models from sED with profiles

In the same way that the sED model with profiles is an extension of the basic and multistage sED models, we would like $\mathcal{R}_0^d$ and $\mathcal{R}_0^s$ to be an extension of the $\mathcal{R}_0$ for these two models. To corroborate this, we consider the special case when we only have one profile, let us say $\mathcal{P} = \{v_1\}$, representing a homogeneous population. In this case, the basic reproductive number for the multistage sED with static profiles, defined in equation (3.6), simplifies to

$$\mathcal{R}_0^{s,v_1} = \frac{1}{1/\mu_R^{(v_1,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v_1,0)}} \beta^{v_1} \kappa_0^{v_1}(0) \frac{\lfloor \varepsilon(0)S(0) \rfloor}{T(0)}, \tag{3.8}$$

where $\varepsilon(0) = \varepsilon^{v_1}(0)$ and $S_{v_1}(0) = S(0)$ as the profile v_1 represents a whole, homogeneous population. On the other hand, the basic reproductive number for the multistage sED with dynamic profiles, defined in equation (3.7), boils down to

$$\mathcal{R}_0^{d,v_1} = \frac{1}{1/\mu_R^{(v_1,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v_1,0)}} \beta_0^{v_1} \kappa_0^{v_1}(0) \frac{\lfloor \varepsilon(0)S(0) \rfloor}{T(0)}, \tag{3.9}$$

where $\varepsilon(0) = \varepsilon^{v_1}(0)$ and $S_{v_1}(0) = S(0)$ as the profile v_1 represents a whole, homogeneous population. In both cases, we recover the expression of the basic reproductive number for the multistage sED (Giral-Barajas et al., 2023). Once again, the only difference in the expressions obtained in both cases lies in the probability of infection given an infectious contact. In this simplified case, if β^{v_1} for the static profiles is the same as $\beta_0^{v_1}$ for the dynamic profiles (both equivalent to β), we obtain the same value of $\mathcal{R}_0^{v_1}$, given by

$$\mathcal{R}_0^{v_1} = \frac{1}{1/\mu_R^{(v_1,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v_1,0)}} \beta \kappa_0^{v_1}(0) \frac{\lfloor \varepsilon(0) S(0) \rfloor}{T(0)}.$$

Furthermore, we would like the $\mathcal{R}_0$ of the sED model with profiles not only to recover the $\mathcal{R}_0$ of the basic sED and the multistage sED, but also to recover the $\mathcal{R}_0$ of SIR-type models. Taking this into account, let us consider the case with one profile, totally exposed susceptible population, and totally exposed infected population ($\kappa_0^{v_1}(0) = \varepsilon^{v_1}(0)$=1). Since $\mu_R^{(v_1,0)}$ is the average time within which an individual clears the pathogen at the first stage of infection and recovers, then $1/\mu_R^{(v_1,0)}$ is the mean rate at which newly infected people recover every time unit, given by γ in the SIR model. Therefore, if we have more than one stage of infection ($k > 1$), equations (3.8) and (3.9) respectively become

$$\mathcal{R}_0^{s,v_1} = \frac{S(0)}{T(0)} \frac{\beta^{v_1}}{\gamma} \quad \text{and} \quad \mathcal{R}_0^{d,v_1} = \frac{S(0)}{T(0)} \frac{\beta_0^{v_1}}{\gamma}, \tag{3.10}$$

and we recover the usual $\mathcal{R}_0$ for the deterministic SIR (Brauer and Castillo-Chavez, 2012), since $S(0)/T(0) \approx 1$ and

$$\mathcal{R}_0 = \frac{\beta}{\gamma}. \tag{3.11}$$

As observed in the previous cases, if $\beta^{v_1} = \beta_0^{v_1} = \beta$ then we have equality between equations (3.10) and (3.11). Moreover, if we have a unique stage of infection ($k = 1$), the mean rate at which infected people die each time unit is $1/\mu_R^{(v_1,0)}$, given by δ in the SIRD model. Therefore, if we have one stage of infection, totally exposed susceptible population, and totally exposed infected population, equations (3.8) and (3.9) respectively take the form of

$$\mathcal{R}_0^{s,v_1} = \frac{S(0)}{T(0)} \frac{\beta^{v_1}}{\gamma + \delta} \quad \text{and} \quad \mathcal{R}_0^{d,v_1} = \frac{S(0)}{T(0)} \frac{\beta_0^{v_1}}{\gamma + \delta}. \tag{3.12}$$

That being the case, we recover the usual $\mathcal{R}_0$ for the deterministic SIRD (Shringi et al., 2021) since $S(0)/T(0) \approx 1$ and

$$\mathcal{R}_0 = \frac{\beta}{\gamma + \delta}. \tag{3.13}$$

Finally, given that $\beta^{v_1} = \beta_0^{v_1} = \beta$ we have equality between equations (3.12) and (3.13).

3.2 $\mathcal{R}_0$ as a weighted average

As demonstrated in the last subsection, the total exposure of susceptible and infected individuals represents a special case in terms of expressions obtained for the general basic reproductive number. On that account, in this section, we explore the total exposure of susceptible and infected individuals with several profiles ($\varepsilon^v(0) = \kappa_0^v(0) = 1$ for all $v \in \mathcal{P}$). In this case, the basic reproductive number with static profiles (3.6) collapses to

$$\mathcal{R}_0^s = \frac{S(0)}{T(0)} \frac{\sum_{v\in\mathcal{P}} \pi(v)\beta^v}{\sum_{v\in\mathcal{P}} \pi(v) \left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)}. \tag{3.14}$$

On the other hand, the basic reproductive number with dynamic profiles (3.7) simplifies to

$$\mathcal{R}_0^d = \frac{S(0)}{T(0)} \frac{\sum_{v\in\mathcal{P}} \pi(v)\beta_0^v}{\sum_{v\in\mathcal{P}} \pi(v) \left(1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}\right)}. \tag{3.15}$$

These expressions are obtained by considering $S^v(0) = \pi^v(0)S(0)$ for equation (3.14) and $I_0^v(0) = \pi^v(0)I_0(0)$ for equation (3.15). The only difference between them lies in the probabilities of infection given an infectious contact β^v and β_0^v. Both expressions show that the rates of change in the infected population depend on the weighted average of the probabilities of infection and the mean rates of recovery and death at group level, as observed for several models.

Now, assuming that there is at least one infectious individual from each profile at the beginning of the epidemic ($I_0^v(0) \geq 1$ for every $v \in \mathcal{P}$), it follows directly from equation (3.6) that in the case of static profiles

$$\mathcal{R}_0^s \leq \frac{\sum_{v\in\mathcal{P}} \lfloor \varepsilon^v(0)S^v(0) \rfloor \beta^v \sum_{v'\in\mathcal{P}} \kappa_0^{v'}(0) \frac{I_0^{v'}(0)}{T(0)}}{\sum_{v\in\mathcal{P}} 1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}}.$$

On the other hand, it follows from equation (3.7) that in the case of dynamic profiles

$$\mathcal{R}_0^d \leq \frac{\sum_{v\in\mathcal{P}} \lfloor \varepsilon(0)S(0) \rfloor \beta_0^v \kappa_0^v(0) \frac{I_0^v(0)}{T(0)}}{\sum_{v\in\mathcal{P}} 1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}}.$$

Therefore, in the case of total exposure of susceptible and infected individuals with several

profiles, we have that for static profiles

$$\mathcal{R}_0^s \leq \frac{S(0)}{T(0)} \sum_{v \in \mathcal{P}} \pi(v) \frac{\beta^v}{1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}} = \sum_{v \in \mathcal{P}} \pi(v) \mathcal{R}_0^{s,v},$$

where $R_0^{s,v}$ is the basic reproductive number for multistage sED with one static profile given by equation (3.8) at the end of subsection 3.1. On the other hand, in the same case with dynamic profiles, we have that

$$\mathcal{R}_0^d \leq \frac{S(0)}{T(0)} \sum_{v \in \mathcal{P}} \pi(v) \frac{\beta_0^v}{1/\mu_R^{(v,0)} + \mathbb{1}_{\{k=1\}}/\mu_I^{(v,0)}} = \sum_{v \in \mathcal{P}} \pi(v) \mathcal{R}_0^{d,v},$$

where $R_0^{d,v}$ is the basic reproductive number for multistage sED with one dynamic profile given by equation (3.9) at the end of subsection 3.1.

Hence, the basic reproductive number in the case of the total exposure of susceptible and infected individuals is bounded from above by the weighted average of the basic reproductive numbers of each profile. This upper bound describes the basic reproductive number for different deterministic and stochastic epidemiological models with heterogeneous populations (Ellison, 2020; Saldaña and Velasco-Hernández, 2022). This bound allows us to compare the basic reproductive number obtained for the multistage sED with profiles, with the basic reproductive number obtained for previous models with heterogeneous populations. In particular, we know that if in our model there is an epidemic outbreak (i.e. $\mathcal{R}_0 > 1$), then there will also be one in the other models. On the other hand, the case in which the $\mathcal{R}_0$ of the multistage sED with profiles is less than one, but the weighted average of the $\mathcal{R}_0^v$ of each profile is greater than one constitutes an interesting case. On top of that, the model allows us to find groups most vulnerable to an infectious disease regardless of whether it will end in an epidemic and thus focus prevention efforts on them.

Given that in this case $\sum_{v \in \mathcal{P}} \pi(v) \mathcal{R}_0^v > 1$, there exists a profile, let us say v^*, for which $\mathcal{R}_0^{v^*} > 1$. In this subpopulation, there will be an epidemic outbreak. However, as the overall $\mathcal{R}_0 < 1$, this outbreak will not affect the whole population. This suggests that there could be epidemic outbreaks within specific groups (those more vulnerable to infection) while not becoming epidemic outbreaks in the whole population. This analysis provides valuable information, as it allows us to study diseases that may go unnoticed for general public while affecting higher-risk groups.

4 Applications: asymptomatic and symptomatic COVID-19

COVID-19 symptoms can range from very mild to severe. Furthermore, some people may have no symptoms but can spread it nevertheless (*asymptomatic transmission*). The number of asymptomatic cases may be the leading underlying cause of large epidemic outbreaks (Gerba, 2009). During COVID-19 pandemic, several variants of SARS-CoV-2 exhibited radically different behaviours, like Delta (B.1.617.2) and Omicron (B.1.1.529). When Delta dominated (July 2021 to November 2021), studies estimated around 40% of unreported infections, while Omicron (December 2021 to March 2022), was associated with 70% unreported infections (Sigal et al., 2022). Experts believe that asymptomatic contagions may have been an essential source of transmission during COVID-19 pandemic (Day, 2020). In the eyes of Delta and Omicron variants, and their proportions of asymptomatic people, this makes sense since Omicron was transmitted more easily than Delta. Regardless of the increase in cases, the case-fatality ratio for Omicron was fewer than for Delta, with estimates of 0.70% for Omicron, compared to 2.01% for Delta (Xia et al., 2024).

To address this issue, we use the multistage model with static profiles (subsection 2.1.1) to represent asymptomatic and symptomatic individuals. We will define the profile $v = 0$ if the individual is asymptomatic and $v = 1$ if the individual presents symptoms when infected. In general, in the sED modelling scheme, it is assumed that a person may recover from any infectious state, and disease cannot reach a certain degree of severity without previously visiting all disease stages of less severity. Here, we allow an infected and symptomatic individual, with profile $v = 1$, to move through all severity stages, according to the multinomial sampling defined in subsection 2.2. In contrast with that, an infected but asymptomatic individual, with profile $v = 0$, will only be allowed to progress to an infectious state corresponding to mild or no symptoms. Note that this implies that an asymptomatic individual never dies from this infection. Furthermore, the infected asymptomatic individuals will have higher exposure than the infected symptomatic ones, as they could be unaware of the infection.

This model can be formulated in terms of the routing probabilities and the mean time spent in each infection stage, $(r^v_{I_i}, r^v_{R_i}, \bar{\tau}^v_{I_i})$ (Giral-Barajas et al., 2023). We can rewrite the

probabilities of moving to the next infectious class or recovering in equation (2.9) as

$$p_I^{(v,i)} \approx \frac{h}{\mu_I^{(v,i)}} = \frac{r_{I_i}^v}{\bar{\tau}_{I_i}^v} \quad \text{y} \quad p_R^{(v,i)} \approx \frac{h}{\mu_R^{(v,i)}} = \frac{r_{R_i}^v}{\bar{\tau}_{I_i}^v}. \tag{4.1}$$

Therefore, we begin by adjusting a multistage sED model for Delta and Omicron variants with routing probabilities towards recovery and death, and mean time endued in each infectious stage recovered from (study group, 2023). Observe that because we have not introduced any profiles yet, we are not considering asymptomatic cases explicitly. To simulate the dynamics of the epidemic, we programmed the model presented in section 2 in Python using the data presented in Table 1.

Table 1: Notation and values used for simulations with the multistage sED model for Delta and Omicron variants. (study group, 2023; Tan et al., 2023; Trobajo-Sanmartín et al., 2022).

Symbol	Values	Units	Variant	Description
T_0	1000	persons	—	Initial population size
S_0	990	persons	—	Initial size of susceptible population
I_0	10	persons	—	Initial size of infected population
R_0	0	persons	—	Initial size of recovered population
D_0	0	persons	—	Initial size of population dead due to disease
β^D	0.26	—	Delta	Probability of infection given a contact with an infected person with Delta variant
$r_{I_0}^D$	1	—	Delta	Routing probability for the first infectious class
$r_{I_1}^D$	0.63	—	Delta	Routing probability for the second infectious class
$r_{I_2}^D$	0.23	—	Delta	Routing probability for the third infectious class
$r_{I_3}^D$	0.16	—	Delta	Routing probability for the fourth infectious class
$\bar{\tau}_{I_0}^D$	9	days	Delta	Mean time elapsed in first infectious class
$\bar{\tau}_{I_1}^D$	2	days	Delta	Mean time elapsed in second infectious class
$\bar{\tau}_{I_2}^D$	4	days	Delta	Mean time elapsed in third infectious class
$\bar{\tau}_{I_3}^D$	13	days	Delta	Mean time elapsed in fourth infectious class
β^O	0.26	—	Omicron	Probability of infection given a contact with an infected person with Omicron variant
$r_{I_0}^O$	1	—	Omicron	Routing probability for the first infectious class
$r_{I_1}^O$	0.502	—	Omicron	Routing probability for the second infectious class
$r_{I_2}^O$	0.12	—	Omicron	Routing probability for the third infectious class
$r_{I_3}^O$	0.16	—	Omicron	Routing probability for the fourth infectious class
$\bar{\tau}_{I_0}^O$	7	days	Omicron	Mean time elapsed in first infectious class
$\bar{\tau}_{I_1}^O$	2	days	Omicron	Mean time elapsed in second infectious class
$\bar{\tau}_{I_2}^O$	5	days	Omicron	Mean time elapsed in third infectious class
$\bar{\tau}_{I_3}^O$	8.5	days	Omicron	Mean time elapsed in fourth infectious class

Simulations without asymptomatic infections using the multistage sED already present low CFR values (Fig. 4.2, panels C and F), as observed for the Delta variant. However, the CFR for Omicron variant turns out to be higher than expected. On top of that, the infection curves for Omicron and Delta are very similar (Fig. 4.2, panels B and E), while the Omicron curve should have a larger peak than the Delta one.

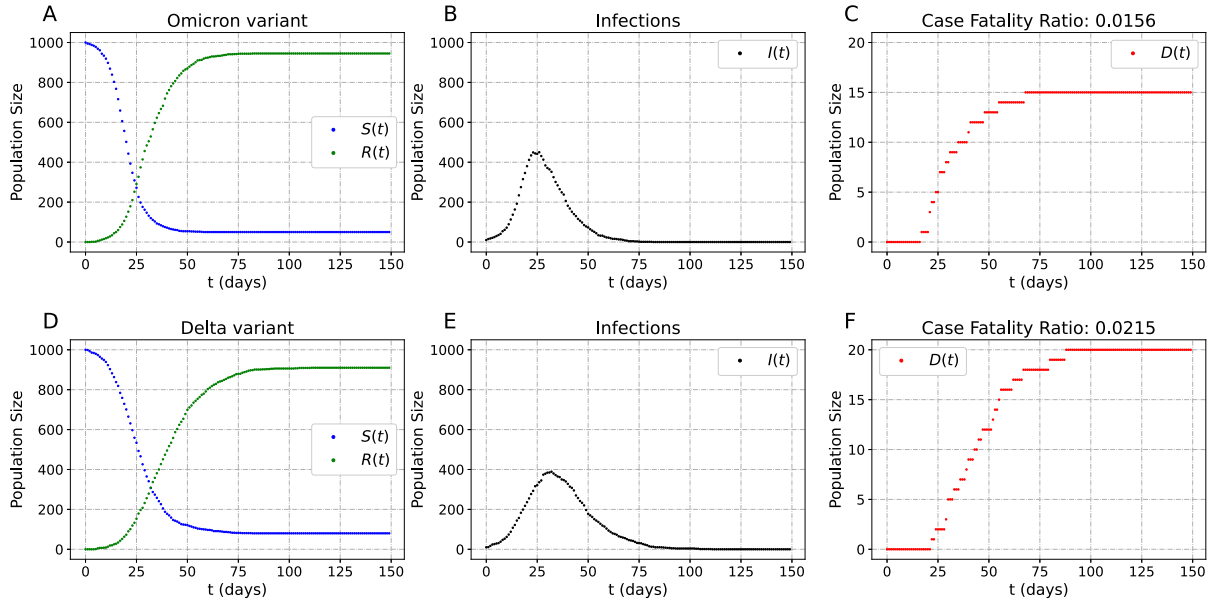

Figure 4.2: Dynamics of the multistage sED adjusted for Omicron (A-C) and Delta (D-F) without asymptomatic cases. The routing probabilities are $r_{I_0} = 1$, $r_{I_1} = 0.502$, $r_{I_2} = 0.12$, $r_{I_3} = 0.16$, and the mean times spent in each infectious stage are $\bar{\tau}_{I_0} = 7$, $\bar{\tau}_{I_1} = 2$, $\bar{\tau}_{I_2} = 5$, $\bar{\tau}_{I_3} = 8.5$ while for the Delta variant $r_{I_0} = 1$, $r_{I_1} = 0.63$, $r_{I_2} = 0.23$, $r_{I_3} = 0.16$, and $\bar{\tau}_{I_0} = 9$, $\bar{\tau}_{I_1} = 2$, $\bar{\tau}_{I_2} = 4$, $\bar{\tau}_{I_3} = 13$ (see Table 1). Furthermore the exposure in both cases is $\kappa^0 = 1$, $\kappa^1 = 0.6$, $\kappa^2 = 0.3$, and $\kappa^3 = 0.1$. Finally, the probability of infection given an infectious contact was $\beta = 0.36$ for Omicron and $\beta = 0.26$ for Delta.

To better describe the dynamics of each variant, we introduce a profile for asymptomatic infections that complements the original profile for symptomatic infections, as mentioned before. In this case, we let asymptomatic individuals progress until second infectious stage before recovering from disease. Following the observed proportion of unreported infections, we will model the population as 70% asymptomatic and 30% symptomatic for Omicron ($\pi(0) = 0.7$ and $\pi(1) = 0.3$) and 40% asymptomatic and 60% symptomatic for Delta ($\pi(0) = 0.4$ and $\pi(1) = 0.6$).

Simulations considering the impact of asymptomatic transmission obtain CFRs closer to those observed for variants for which the asymptomatic proportion is higher, such as Omicron, and maintain the already well-fitted CFRs for variants with lower asymptomatic proportion, such as Delta (Fig. 4.3, panels C and F). Furthermore, in the simulations for

Omicron, the model predicts a higher number of infected people, of which the majority are asymptomatic (Fig. 4.3, panel B green curve). In this way, the multistage sED with asymptomatic and symptomatic profiles, allows us to reproduce dynamics similar to those observed throughout the COVID-19 pandemic for the Delta and Omicron variants.

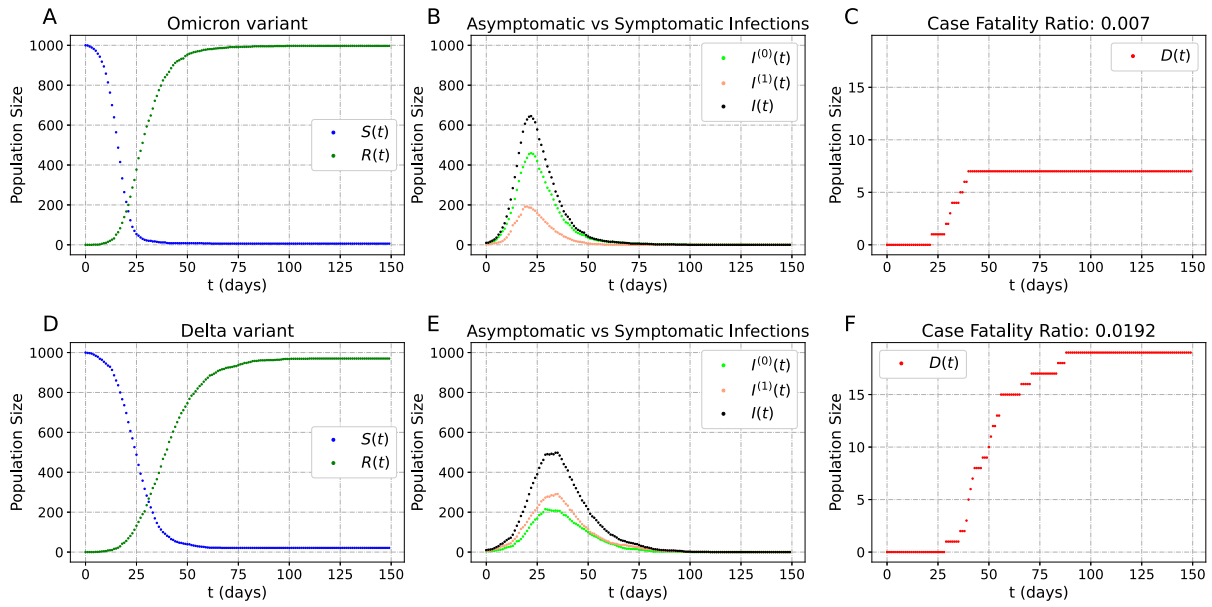

Figure 4.3: Dynamics of the multistage sED with profiles adjusted for Omicron (A-C) and Delta (D-F) with symptomatic and asymptomatic cases. The routing probabilities and mean times for symptomatic cases for both variants are the same as Figure 4.2. For the Omicron variant $\pi(0) = 0.7$ and $\pi(1) = 0.3$, and the routing probabilities and mean times for asymptomatic are $r^0_{I_0} = 1$, $r^0_{I_1} = r^0_{I_2} = r^0_{I_3} = 0$, and $\bar{\tau}^0_{I_0} = \bar{\tau}^0_{I_1} = 7$, $\bar{\tau}^0_{I_2} = \bar{\tau}^0_{I_3} = \infty$. For the Delta variant $\pi(0) = 0.4$ and $\pi(1) = 0.6$, and the routing probabilities and mean times for asymptomatic are $r^0_{I_0} = 1$, $r^0_{I_1} = r^0_{I_2} = r^0_{I_3} = 0$, and $\bar{\tau}^0_{I_0} = \bar{\tau}^0_{I_1} = 9$, $\bar{\tau}^0_{I_2} = \bar{\tau}^0_{I_3} = \infty$. Furthermore the symptomatic exposition for both cases is $\kappa^0 = 1$, $\kappa^1 = 0.6$, $\kappa^2 = 0.3$, and $\kappa^3 = 0.1$, and the asymptomatic exposition for both cases is $\kappa^0 = \kappa^1 = \kappa^2 = \kappa^3 = 1$. Finally, the probability of infection given an infectious contact was $\beta = 0.36$ for Omicron and $\beta = 0.26$ for Delta.

Despite being a simple example of the use of profiles, it demonstrates the power of this model. We obtain considerably different dynamics from the multistage sED by introducing a single input profile with only two possible values. The first notable result is the increase in the infected curve for Omicron, for which more than half of the population gets to be infectious at one point, and reproducing the lower case-fatality ratio. The second notable result is how, when considering different asymptomatic proportions for the Delta and Omicron variants with a difference of less than double, the resulting incidence in CFR is more than three times.

References

Linda J.S. Allen. *An introduction to stochastic processes with applications to biology*. CRC press, 2010.

Fred Brauer and Carlos Castillo-Chavez. *Mathematical models in population biology and epidemiology*. Springer, 2012.

Tom Britton, Etienne Pardoux, Franck Ball, Catherine Laredo, David Sirl, and Viet Chi Tran. *Stochastic epidemic models with inference*, volume 2255. Springer, 2019.

Michael Day. COVID-19: four fifths of cases are asymptomatic, China figures indicate, 2020.

Glenn Ellison. Implications of heterogeneous SIR models for analyses of COVID-19. Technical report, National Bureau of Economic Research, 2020.

Charles P. Gerba. Environmentally transmitted pathogens. In *Environmental microbiology*, pages 445–484. Elsevier, 2009.

José Giral-Barajas, Carlos Ignacio Herrera-Nolasco, Marco Arieli Herrera-Valdez, and Sergio I. López. A probabilistic approach for the study of epidemiological dynamics of infectious diseases: Basic model and properties. *Journal of Theoretical Biology*, 572: 111576, 2023.

Ondrej Hradsky and Arnost Komarek. Demographic and public health characteristics explain large part of variability in COVID-19 mortality across countries. *European Journal of Public Health*, 31(1):12–16, 01 2021. ISSN 1101-1262. doi: 10.1093/eurpub/ckaa226. URL https://doi.org/10.1093/eurpub/ckaa226.

William Ogilvy Kermack and Anderson G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.

Naveen L. Pereira, Ferhaan Ahmad, Mirnela Byku, Nathan W. Cummins, Alanna A. Morris, Anjali Owens, Sony Tuteja, and Sharon Cresci. COVID-19: understanding inter-individual variability and implications for precision medicine. In *Mayo Clinic Proceedings*, pages 446–463. Elsevier, 2021.

Fernando Saldaña and Jorge X. Velasco-Hernández. Modeling the COVID-19 pandemic: a primer and overview of mathematical epidemiology. *SeMA Journal*, pages 1–27, 2022.

Sakshi Shringi, Harish Sharma, Pushpa Narayan Rathie, Jagdish Chand Bansal, and Atulya Nagar. Modified sird model for COVID-19 spread prediction for northern and southern states of India. *Chaos, Solitons & Fractals*, 148:111039, 2021.

Alex Sigal, Ron Milo, and Waasila Jassat. Estimating disease severity of Omicron and Delta SARS-CoV-2 infections. *Nature Reviews Immunology*, 22(5):267–269, 2022.

P.D.N.N. Sirisena, Shakuntala Mahilkar, Chetan Sharma, Jaspreet Jain, and Sujatha Sunil. Concurrent dengue infections: Epidemiology & clinical implications. *The Indian Journal of Medical Research*, 154(5):669, 2021.

COVID-19 Omicron Delta study group. Clinical progression, disease severity, and mortality among adults hospitalized with covid-19 caused by the omicron and delta sars-cov-2 variants: A population-based, matched cohort study. *PloS one*, 18:e0282806, 2023. ISSN 1932-6203. doi: 10.1371/JOURNAL.PONE.0282806. URL https://pubmed.ncbi.nlm.nih.gov/37104488/.

Sophia T. Tan, Ada T. Kwan, Isabel Rodríguez-Barraquer, Benjamin J. Singer, Hailey J. Park, Joseph A. Lewnard, David Sears, and Nathan C. Lo. Infectiousness of sars-cov-2 breakthrough infections and reinfections during the omicron wave. *Nature Medicine 2023 29:2*, 29:358–365, 1 2023. ISSN 1546-170X. doi: 10.1038/s41591-022-02138-x. URL https://www.nature.com/articles/s41591-022-02138-x.

James M. Trauer, Peter J. Dodd, M. Gabriela M. Gomes, Gabriela B. Gomez, Rein M.G.J. Houben, Emma S. McBryde, Yayehirad A. Melsew, Nicolas A. Menzies, Nimalan Arinaminpathy, Sourya Shrestha, et al. The importance of heterogeneity to the epidemiology of tuberculosis. *Clinical Infectious Diseases*, 69(1):159–166, 2019.

Camino Trobajo-Sanmartín, Iván Martínez-Baz, Ana Miqueleiz, Miguel Fernández-Huerta, Cristina Burgui, Itziar Casado, Fernando Baigorría, Ana Navascués, Jesús Castilla, and Carmen Ezpeleta. Differences in transmission between sars-cov-2 alpha (b.1.1.7) and delta (b.1.617.2) variants. *Microbiology Spectrum*, 10, 4 2022. ISSN 21650497. doi: 10.1128/SPECTRUM.00008-22. URL /pmc/articles/PMC9045255//pmc/articles/PMC9045255/?report=abstracthttps://www.ncbi.nlm.nih.gov/pmc/articles/PMC9045255/.

Qianhang Xia, Yujie Yang, Fengling Wang, Zhongyue Huang, Wuqi Qiu, and Ayan Mao. Case fatality rates of COVID-19 during epidemic periods of variants of concern: A meta-analysis by continents. *International Journal of Infectious Diseases*, 141:106950, 2024.

Yanjie Zhao, Roujian Lu, Jun Shen, Zhengde Xie, Gaoshan Liu, and Wenjie Tan. Comparison of viral and epidemiological profiles of hospitalized children with severe acute respiratory infection in Beijing and Shanghai, China. *BMC Infectious Diseases*, 19(1): 1–8, 2019.

Fei Zhou, Ting Yu, Ronghui Du, Guohui Fan, Ying Liu, Zhibo Liu, Jie Xiang, Yeming Wang, Bin Song, Xiaoying Gu, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet*, 395(10229):1054–1062, 2020.