Generic Approach For Mathematical Model of Multi-Strain Pandemics

Teddy Lazebnik[1] *, Svetlana Bunimovich-Mendrazitsky[2], with the Lorem Ipsum Consortium[¶]

1 Department of Cancer Biology, Cancer Institute, University College London, UK
2 Department of Mathematics, Ariel University, Israel

* Corresponding author: t.lazebnik@ucl.ac.uk

Abstract

Pandemics with multi-strain have become a major concern. We introduce a new model for assessing the connection between multi-strain pandemic and the mortality rate, basic reproduction number, and the maximum of infected individuals. The proposed model provides a general mathematical approach for representing multi-strain pandemics, generalizing for an arbitrary number of strains. We show the proposed model fits well with epidemiological historical data world health over a long period. From a theoretical point of view, we show that the increasing number of strains increases logarithmically the maximum number of infected individuals and the mean mortality rate. Moreover, the mean basic reproduction number is statistically identical to the single, most aggressive strain pandemic for multi-strain pandemics.

1 Introduction and Related Work

Humanity has experienced multiple types of disasters over the centuries [1–4]. One of them is (local and global) pandemics that cause significant mortality [5]. Moreover, recent studies show that the occurrence rate of new pandemics has increased in the last century, resulting in an increased number of pandemics and their influence [6]. Some of these pandemics exert a global influence such as HIV/AIDS that killed 680 thousand individuals only in 2020 according to the World Health Organization (WHO)[1] or the COVID-19 pandemic that killed 4.5 million individuals and infected around 440 million individuals worldwide during its first 18-months [7]. As a result, the need for policymakers to be able to control a pandemic spread is becoming more relevant by the day [8].

Moreover, due to multiple socioeconomic processes, there is an increase in the speed at which new infections are spread [9]. To be exact, globalization has facilitated strain spread among countries through the growth of trade and travel [10]. Diseases are usually caused by pathogenic agents, including viruses and bacteria, which can be denoted as multiple variants, generally named strains. The presence of multi-strain for a pathogen imposes a new challenge to control the spread of disease [11]. Since new strains occur as it reproduces in new hosts, the large population of infected individuals offers a fertile ground for new strains to appear [12, 13]. For example, in the case of

[1]The full report is available online at: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/data-on-the-size-of-the-hiv-aids-epidemic

COVID-19, already in the first year and a half of the pandemic, four (globally common) strains were detected [7].

Most diseases have several pathogenic strains, which can make it difficult to fight the disease and lead to rich dynamics. However, their dynamic properties have not been adequately studied [11]. Hence, a better understanding of future pandemics with several strains is a necessary step to ensure the ability of the global community in handling the next pandemic. One approach to tackle this challenge is using epidemiological-mathematical models, which allows us to simulate and investigate multiple scenarios in a safe, cheap, and manageable environment. A large portion of these epidemiological models are based on the Susceptible-Infectious-Removed (*SIR*) model [14]. Over the years, researchers have introduced different extensions to the *SIR* model in order to obtain a more accurate model for biological [15], economic [16, 17] spatial [18–20], and pandemic management [8, 21, 22] properties of a particular disease or socio-epidemiological scenario. These extensions are natural as the *SIR* disease transmission model is derived assuming multiple strong assumptions. For example, the *SIR* model assumes that the population is large and dense or that the infection rate is constant [14]. The authors extend this basic model in many directions by relaxing some assumptions. As such, the mathematical analysis quickly becomes significantly more sophisticated [23].

Cooper et al. [24] used the *SIR* model on the COVID-19 pandemic while relaxing the assumption that the population is mixing homogeneously and that the total population is constant in time. The authors show that the model has a fair fitting on six countries (China, South Korea, India, Australia, USA, Italy). However, the model failed to capture sharp changes in the dynamics due to the pandemic modifications on governance intervention policies.

Another extension of the *SIR* model for the Polio pandemic is proposed by Agarwal and Bhadauria [25]. The authors introduced the fourth stage - vaccinated individuals, resulting in a *SIR*V model. The numerical simulation of the model results in a promising outcome. Nonetheless, the evaluation is limited to a small size (up to a few hundred individuals), and the generalization to larger populations can be less accurate due to the increased chance that a strain occurs during the pandemic and changes its dynamics [26].

Similarly, Bunimovich-Mendrazitsky and Stone [27] proposed a two-age group, adult and children, for the Polio pandemic spread. Using the model in [27], the extraordinary jump in the number of paralytic polio cases that emerged at the beginning of the 20th century can be explained. The model does not take into consideration some strains of Polio [28] which results in an increased divergence from the actual dynamics over time.

In addition, one of the main extensions of the *SIR* model is the *SIRD* (D-Dead) model, as this model takes is able to represent the reinfection process and the death of individuals due to the pandemic [29–31]. This model better represents the biological-clinical dynamics in human populations as the long-term immunity memory is reduced over time making the individual susceptible again [32, 33]. We based our model on this extension as it allows reinfection in several strains of the original strain.

The mentioned models and other models that extend the *SIR* model can fairly fit and predict the course of a pandemic spread [34, 35]. However, the models are not fitted to capture sharp changes in the dynamics due to the pandemic modifications. One reason is the lack of modelization of pandemics with multi-strain.

Indeed, the occurrence of pandemics with multiple mutations is common. For example, Minayev and Ferguson [36] investigate the interaction between epidemiological and evolutionary dynamics for antigenically variable pathogens. The authors proposed a set of relatively simple deterministic models of the transmission dynamics of multi-strain pathogens which give increased biological realism. However, these models

assume clinical-epidemiological dynamics that hold only for a subset of pathogens with cross-immunity of less than 0.4 [36]. In a similar manner, Dang et al. [37] developed a multi-scale immuno-epidemiological model of influenza viruses including direct and environmental transmission. The authors showed how two time-since-infection structural variables outperform classical SIR models of influenza. During the modelization, they used a within-host model that holds only for the influenza pandemic. In addition, Gordo et al. [12] proposed a *SIRS* model with reinfection and selection with two strains. The authors used a metapopulation of individuals where each individual is depicted as a vector in the metapopulation. This model has been validated on the influenza pandemic in the State of New York (USA), based on the genetic diversity of influenza gathered between 1993 and 2006, showing supreme results compared to other *SIR*-based models [12]. Nonetheless, the sophistication of the model is both in its strength and shortcoming, from an analytical point of view, due to its stochastic and chaotic nature.

Moreover, the usage of multi-strain models that are used for specific pathogens is not restricted to influenza. Marquioni and de Aguiar [38] proposed a model where a pandemic starts with a single strain and the other strains occur in a stochastic manner as a by-product of the infection. The authors fitted their model onto the COVID-19 pandemic in Chine showing improved results when strain dynamics are taken into consideration compared to the other case [38]. Likewise, Khayar and Allali [39] proposed a *SEIR* (E-exposed) model for the COVID-19 pandemic with two strains. The authors analyzed the influence of the delay between exposure and becoming infectious on several epidemiological properties. Furthermore, they proposed an extension to the model (in the appendix of the manuscript) for multi-strain dynamics. In their model, an individual can be infected only once and develop immunity to all strains [39]. In our model, we relax this assumption, allowing individuals to be infected once by each strain. Comparably, Gubar et al. [40] proposed an extended *SIR* model with two strains with different infection and recovery rates. The authors considered a group of latent individuals who are already infected but do not have any clinical symptoms.

Correspondingly, Aleta et al. [41] extended the SIRS model on a metapopulation where individuals are distributed in sub-populations connected via a network of mobility flows. They show that spatial fragmentation and mobility play a key role in the persistence of the disease whose maximum is reached at intermediate mobility values. Their model assumes a fixed number of locations (using a graph-based model) such that each location has a unique strain-like simulation. Furthermore, Di Giamberardino et al. [42] proposed a multi-group model formed by interconnected SEIR-like structures which include asymptomatic infected individuals. The authors fitted the data to the COVID-19 pandemic in Italy to study the influence of different IPs on the pandemic spread. The interconnection between the groups in the model is represented by the mobility of individuals between them. The model represents somewhat multi-strains as each group has different epidemiological parameter values and the transformation between them. However, the authors do not handle the case where an individual has been infected by one strain and later infected by others which are known from multiple clinical and biological studies [43–46].

In this research, we developed an extension of the *SIRD*-based model which allows an arbitrary number of strains $|M|$ that originated from a single strain and is generic for any type of pathogen. The model allows each strain to have its unique epidemiological properties. In addition, we developed a computer simulation that provides an *in silico* tool for evaluating several epidemiological properties such as the mortality rate, max infections, and average basic reproduction number of a pandemic. The proposed model allows for a more accurate investigation of the epidemiological dynamics while keeping the data required to use the model relatively low.

This paper is organized as follows: In Section 2, we introduce our multi-strain

epidemiological model. Based on the model, we present a numerical analysis of three epidemiological properties as a function of the number of strains ($|M|$). In Section 3, we present the implementation of the model for the case of two strains ($|M| = 2$) and provide an analytical analysis of the stable equilibria states of the model and a basic reproduction number analysis. Afterward, we show the ability of the model to fit historical epidemiological data known to have two strains. In Section 4, we discuss the main advantages and limitations of the model and propose future work.

2 Multi-Strain Model

The multi-strain epidemiological model considers a constant population with a fixed number of individuals N. We assume a pandemic has $M := \{1, \ldots, m\}$ strains. Moreover, two options are possible: a) strains $[2, \ldots m]$ are mutations arising from one pathogen as a result of the mutation process; b). the disease is characterized by the emergence of m pathogenic strains but an individual cannot be infected by more than one strain of the virus at a time.

Each individual belongs to one of the three groups. 1) Infectious with strain $i \in M$ and history of recoveries $J \in P(M)$ (the power set of the strain and its strain set) represented by $R_J I_i$, which maps to the infection (I) state in the *SIRD* model; 2) Recovered with history $J \in P(M)$ represented by R_J, which maps to the recovered (R) state in the *SIRD* model; and 3) Dead (D) such that

$$N = \Sigma_{J \in P(M \setminus \{i\}), i \in M}(R_J I_i(t)) + \Sigma_{J \in P(M)}(R_J(t)) + D(t), \tag{1}$$

where $i \in M$ is the index of a strain and $J \in P(M)$ is the set of strains an individual already suffered from. For example, $R_\emptyset$ is the group of individuals that do not have a recovery history and are susceptible to all $|M|$ strains which is a private case of R_J where $J = \emptyset$, which is isomorphic to the susceptible (S) state in the *SIRD* model. The proposed model for $|M| = 1$ is isomorphic to the *SIRD* model (the proof is provided in the supplementary material, Section S2). A schematic transition between disease stages of an individual is shown in Fig. 1.

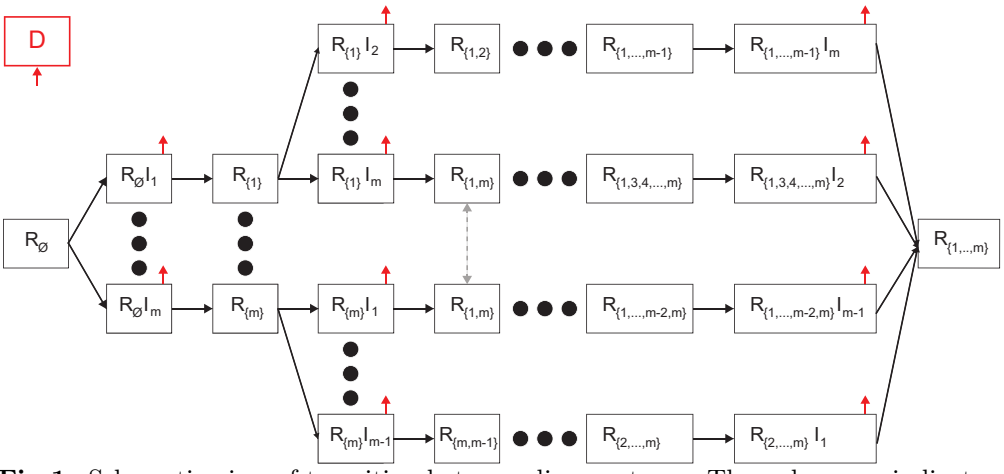

Fig 1. Schematic view of transition between disease stages. The red arrows indicates that individuals from the source stage can transferred to the dead stage. Individuals in $R_J I_i$ stages are necessarily transferred to the respective $R_{J \cup i}$ stages (or dead stage), while individuals in the R_J stages move to $R_J I_l$ stage if they are infected by an individual that is infectious in strain $l \in M$.

Individuals in the Recovered (R_J) group have immunity for the strains $k \in J$ and are susceptible to the infection strains $M \backslash J$. When an individual in this group is exposed to a strain $i \in M \backslash J$, the individual is transferred to the Infectious with history of recoveries group ($R_J I_i$) at a rate $\beta_{J,i}$. The individual stays in this group on average $\gamma_{J,i}$ days, after which the individual is transferred to the Recovered group ($R_{J \cup \{i\}}$) or the Dead group (D). Therefore, at a rate of $(1 - \phi_{J \backslash i})$, of infection by strain i with a history of recoveries from strains J, individuals remain seriously ill or die while others recover. The recovered are again healthy, no longer contagious, and immune from future infection of the same strain. The epidemiological dynamics are described in Eqs. (2-4).

In Eq. (2), $\frac{dR_J I_i(t)}{dt}$ is the dynamical amount of individuals that recovered from a group of strains J and are infected with a strain i over time. It is affected by the following two terms. First, individuals who recovered from group J of strains become infected with strain i, with rate $\beta_{J,i}$. These individuals can be infected by any individual with a strain i who has recovered from any group K of strains so that $i \notin K$. Second, individuals recover from strains $J \cup \{i\}$ with rate $\gamma_{J,i}$. For each strain i, the group i can be any subgroup of the group M, so that $i \notin J$.

$$\frac{dR_J I_i(t)}{dt} = -\gamma_{J,i} R_J I_i(t) + \beta_{J,i} R_J(t) \sum_{K \in P(M), i \notin K} R_K I_i(t). \tag{2}$$

In Eq. (3), $\frac{dR_J(t)}{dt}$ is the dynamical amount of individuals that recovered from a group of strains $J \in P(M)$ over time. It is affected by the following two terms. First, for each strain $i \in J$, an individual who has recovered from group $J \backslash \{i\}$ of strains and is infected with strain i, recovers at rate $\gamma_{J \backslash \{i\},i}$ with probability of $\phi_{J \backslash \{i\},i}$. Second, individuals infected by strain i with rate $\beta_{J,i}$. These individuals can be infected by any individual with a strain i who has recovered from any group K of strains, so that $i \notin K$.

$$\frac{dR_J(t)}{dt} = \sum_{i \in J} \left(\gamma_{J \backslash \{i\},i} \phi_{J \backslash \{i\},i} R_{J \backslash \{i\}} I_i(t) \right) - \sum_{i \in M \backslash J} \left(\beta_{J,i} R_J(t) \sum_{K \in P(M), i \notin K} R_K I_i(t) \right). \tag{3}$$

In Eq. (4), $\frac{dD(t)}{dt}$ is the dynamical amount of dead individuals over time. For each strain i, and for each group $J \backslash \{i\}$, infected individuals that do not recover are dying at rate $\gamma_{J \backslash \{i\},i}$ with the complete probability $(1 - \phi_{J \backslash \{i\},i})$.

$$\frac{dD(t)}{dt} = \sum_{i \in M, J \in P(M)} \gamma_{J \backslash \{i\},i} (1 - \phi_{J \backslash \{i\},i}) R_{J \backslash \{i\}} I_i(t). \tag{4}$$

The dynamics of Eqs. (2-4) are summarized in Eq. (5).

$$\begin{aligned}
&\tfrac{dR_J I_i(t)}{dt} = -\gamma_{J,i} R_J I_i(t) + \beta_{J,i} R_J(t) \textstyle\sum_{K \in P(M), i \notin K} R_K I_i(t), \\
&\tfrac{dR_J(t)}{dt} = \textstyle\sum_{i \in J} \left(\gamma_{J \backslash \{i\},i} \phi_{J \backslash \{i\},i} R_{J \backslash \{i\}} I_i(t) \right) - \sum_{i \in M \backslash J} \left(\beta_{J,i} R_J(t) \sum_{K \in P(M), i \notin K} R_K I_i(t) \right), \\
&\tfrac{dD(t)}{dt} = \textstyle\sum_{i \in M, J \in P(M)} \gamma_{J \backslash \{i\},i} (1 - \phi_{J \backslash \{i\},i}) R_{J \backslash \{i\}} I_i(t),
\end{aligned} \tag{5}$$

The initial conditions of Eq. (5) are defined for the beginning of a pandemic as follows:

$$\begin{aligned}
&R_\emptyset(0) = N - m, \ \forall i \in M : R_\emptyset I_i(0) = 1, \\
&\forall J \in P(M) \backslash \emptyset \wedge i \in M \backslash J : R_J(0) = R_J I_i(0) = 0, \ D(0) = 0.
\end{aligned} \tag{6}$$

2.1 Epidemiological properties

Based on the proposed model, and since for all the cases where $|M| > 2$ it is extremely hard (or even impossible) to obtain an analytical result, we evaluated three important epidemiological properties to see the influence of the number of strains on the pandemic spread: mean basic reproduction number, mortality rate, and a maximum number of the infectious. Formally, these properties can be defined as follows.

First, the *mean basic reproduction number* is the mean basic reproduction number overtime during the course of the pandemic. Therefore, it takes the form:

$$E[R_0(t)] := E[\forall J \in P(M) : \Sigma_{i \in M}(\frac{R_J I_i(t+1) - R_J I_i(t)}{R_{J,i}(t+1) - R_{J,i}(t)})].$$

Second, the *mortality rate* is defined as the number of deaths due to the pandemic divided by the number of infections at some period of time. If not stated otherwise, we assume the mortality rate refers to the entire duration of the pandemic. Hence, it takes the form:

$$\text{mortality rate}(t_0, t_1) := \frac{D(t_1) - D(t_0)}{\Sigma_{J \in P(M)} |J| * (R_J(t_1) - R_J(t_0)}.$$

Finally, the *maximum number of infectious* refers to the cumulative number of infections that occur during the pandemic. Thus, it takes the form:

$$\text{maximum number of infectious}(t_0, t_1) := \max_{t \in [t_0, t_1]} \big(\forall J \in P(M) : \Sigma_{i \in M}(R_J I_i(t))\big).$$

In addition, we define the *most aggressive* strain using the following metric: a strain k is considered more aggressive than strain l if and only if:

$$||[\forall J \in P(M) : (\beta_{J,k}, 1 - \gamma_{J,k}, 1 - \phi_{J,k})]|| > ||[\forall J \in P(M) : (\beta_{J,l}, 1 - \gamma_{J,l}, 1 - \phi_{J,l})]||.$$

2.2 Numerical simulation

Using numerical simulation we aim to study the connection between the number of strains $|M|$ and the proposed epidemiological properties. We numerically solved the model presented in Eq. (5) for the case where $|M| \in [1, \ldots, 10]$ using the fourth-order Runge-Kutta algorithm [47]. The model parameters are generated randomly as follows. The infection rates $\beta_{J,i}$ are uniformly distributed in $[0.01, 0.10]$; the recovery rates $\gamma_{J,i}$ are uniformly distributed in $[0.03, 0.33]$; and the recovery probabilities $\phi_{J,i}$ are uniformly distributed in $[0.90, 0.99]$ for each strain. The ranges were picked to simulate a large space of possible pandemics, without taking into consideration biological properties associated with cross-immunity between strains. In addition, we assume the population size is 10 million individuals to approximate (in order of magnitude) a European metropolitan area. The simulation begins with one person getting infected by each strain. In addition, it is assumed that no individuals have recovered or died as a result of the pandemic. Formally, the initial conditions take the form:

$$R_\emptyset(0) = N - |M|,\ \forall i \in M : R_\emptyset I_i(0) = 1,\ D(0) = 0,\ \forall J \in P(M) \backslash \emptyset, i \in M : R_J I_i = R_J = 0.$$

Moreover, due to the stochastic nature of the simulation originating in the large ranges of values allocated to the model parameters, the simulation is repeated 1000 times, and the mean $\pm$ standard division is presented. Using this generation we compute the connection between $|M|$ and the mean basic reproduction number, max infected individuals, and mean mortality rate.

The *mean basic reproduction number* ($E[R_0]$) has been evaluated for each simulation, divided into two cases: the case where strain has unique parameter values and the case

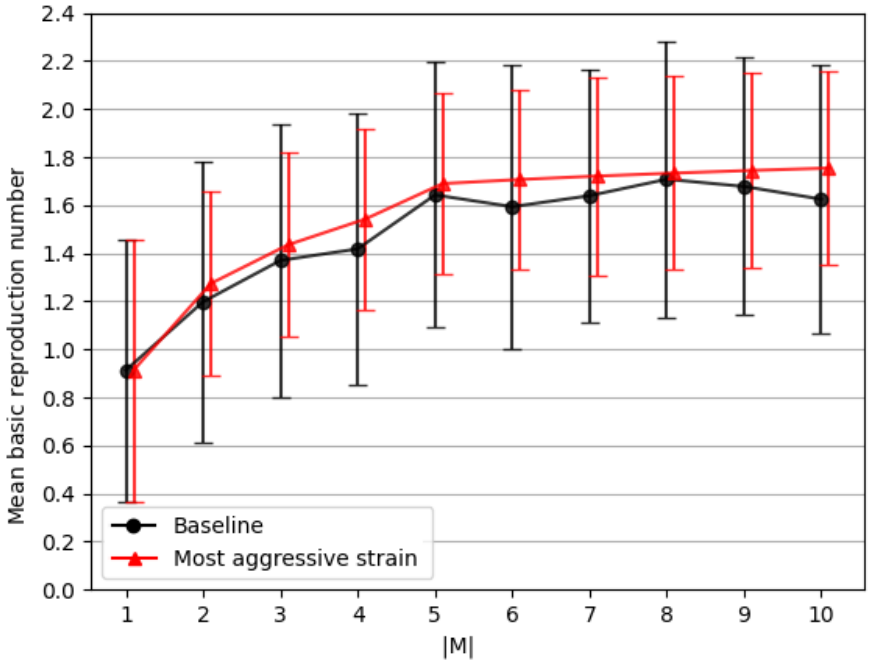

Fig 2. The mean base reproduction number ($E[R_0(t)]$) as a function of the number of strains ($|M|$). The black (with circle markers) line indicates the baseline dynamics of the simulation where each strain has a unique parameter values. On the other hand, the red (with triangle markers) line indicates the case where all the strains parameters values have been replaced with the one of the most aggressive strain.

where the parameters for each strain are replaced with the parameter value of the most aggressive strain, as defined in Section 2.1, as shown in Fig. 2.

The *maximum number of infected individuals* as a function of the number of strains ($|M|$) has been computed and shown in Fig. 3. The solid (black) line with the dots represents the numerically calculated values with one standard deviation. Moreover, the fitting function is calculated using the least mean square (LMS) method [48] and shown as the dashed (blue) line. In order to use the LMS method, one needs to define the family function approximating a function. The family function that has been chosen is $f(m) = p_1 log(m) + p_2$, resulting in

$$E[R_0](m) = 0.103 log(m) + 0.068. \quad (7)$$

The fitting function was obtained with a coefficient of determination $R^2 = 0.79$.

The *mean mortality rate* as a function of the number of strains has been computed and presented in Fig. 4. Similarly, the dots are the calculated values from the simulator and the dotted line is a fitting function that is computed using the LMS with the family function $f(m) = p_1 log(m) + p_2$, resulting in

$$E[R_0](m) = 0.0341 log(m) + 0.0124. \quad (8)$$

The fitting function was obtained with a coefficient of determination $R^2 = 0.89$.

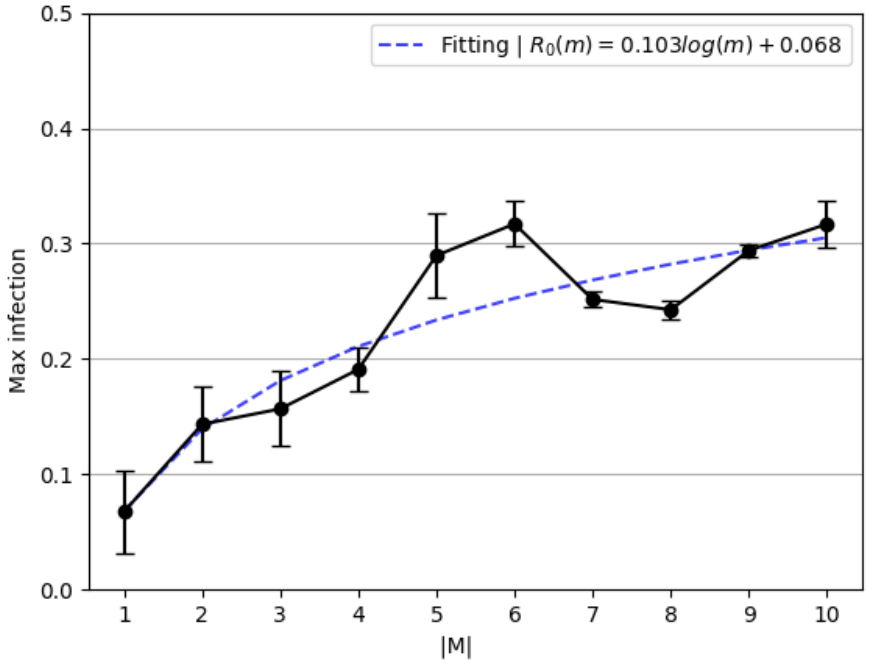

Fig 3. Maximum number of infectious individuals at the same time as a function of the number of strains ($|M|$).

3 Two Strains Model

The two strains epidemiological model considers a constant population with a fixed number of individuals N. We assume a pandemic has two strains $M = \{1, 2\}$. We define a system of eight ordinary differential equations (ODEs) corresponding to eight possible epidemiological states: susceptible ($R_\emptyset$), infected by strain 1 ($R_\emptyset I_1$), infected by strain 2 ($R_\emptyset I_2$), recovered from strain 1 ($R_{\{1\}}$), recovered from strain 2 ($R_{\{2\}}$), recovered from strain 1 and infected by strain 2 ($R_{\{1\}} I_2$), recovered from strain 2 and infected by strain 1 ($R_{\{2\}} I_1$), recovered from both strains ($R_{\{1,2\}}$), and dead (D). The full explanation of how one obtains the model is provided in the supplementary materials, Section S1. A schematic transition between disease stages of an individual for the case of $|M| = 2$ is shown in Fig. 5. Thus, the model for two strains is described by eight equations as follows.

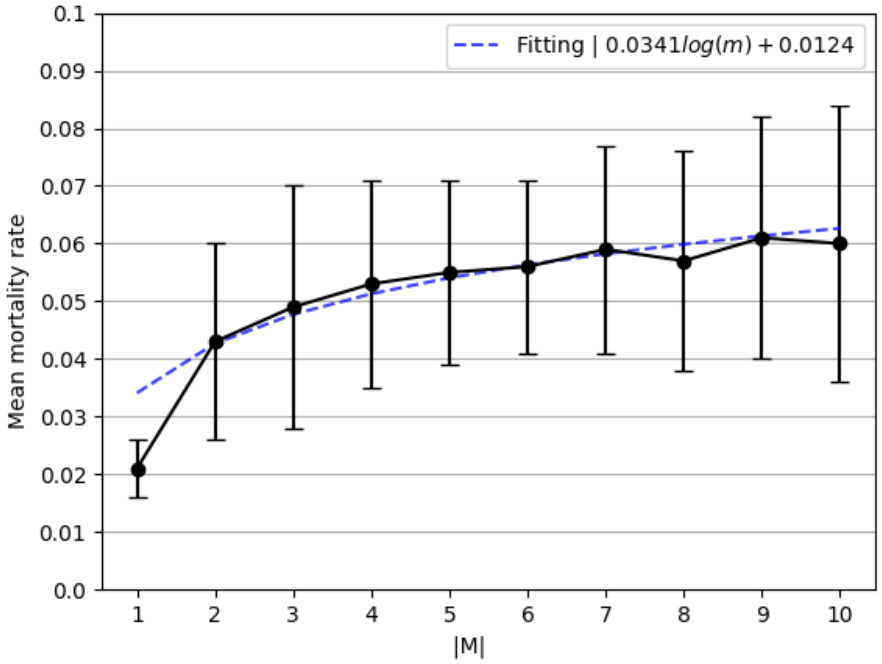

Fig 4. Mortality rate as function of the number of strains ($|M|$).

$$
\begin{aligned}
\tfrac{dR_{\emptyset}I_1(t)}{dt} &= \beta_{\emptyset,1}(R_{\emptyset}I_1(t) + R_{\{2\}}I_1(t))R_{\emptyset}(t) - \gamma_{\emptyset,1}R_{\emptyset}I_1(t),\\
\tfrac{dR_{\{2\}}I_1(t)}{dt} &= \beta_{\{2\},1}(R_{\{2\}}I_1(t) + R_{\emptyset}I_1(t))R_{\{2\}}(t) - \gamma_{\{2\},1}R_{\{2\}}I_1(t),\\
\tfrac{dR_{\emptyset}I_2(t)}{dt} &= \beta_{\emptyset,2}(R_{\emptyset}I_2(t) + R_{\{1\}}I_2(t))R_{\emptyset}(t) - \gamma_{\emptyset,2}R_{\emptyset}I_2(t),\\
\tfrac{dR_{\{1\}}I_2(t)}{dt} &= \beta_{\{1\},2}(R_{\{1\}}I_2(t) + R_{\emptyset}I_2(t))R_{\{1\}}(t) - \gamma_{\{1\},2}R_{\{1\}}I_2(t),\\
\tfrac{dR_{\emptyset}(t)}{dt} &= -R_{\emptyset}(t)\big(\beta_{\emptyset,1}(R_{\emptyset}I_1(t) + R_{\{2\}}I_1(t)) + \beta_{\emptyset,2}(R_{\emptyset}I_2(t) + R_{\{1\}}I_2(t))\big),\\
\tfrac{dR_{\{1\}}(t)}{dt} &= \gamma_{\emptyset,1}\phi_{\emptyset,1}R_{\emptyset}I_1(t) - \beta_{\{1\},2}(R_{\{1\}}I_2(t) + R_{\emptyset}I_2(t))R_{\{1\}}(t),\\
\tfrac{dR_{\{2\}}(t)}{dt} &= \gamma_{\emptyset,2}\phi_{\emptyset,2}R_{\emptyset}I_2(t) - \beta_{\{2\},1}(R_{\{2\}}I_1(t) + R_{\emptyset}I_1(t))R_{\{2\}}(t),\\
\tfrac{dR_{\{1,2\}}(t)}{dt} &= \gamma_{\{2\},1}\phi_{\{2\},1}R_{\{2\}}I_1(t) + \gamma_{\{1\},2}\phi_{\{1\},2}R_{\{1\}}I_2(t),\\
\frac{dD(t)}{dt} &= \gamma_{\emptyset,1}(1-\phi_{\emptyset,1})R_{\emptyset}I_1(t) + \gamma_{\{2\},1}(1-\phi_{\{2\},1})R_{\{2\}}I_1(t)\\
&\quad +\gamma_{\emptyset,2}(1-\phi_{\emptyset,2})R_{\emptyset}I_2(t) + \gamma_{\{1\},2}(1-\phi_{\{1\},2})R_{\{1\}}I_2(t).
\end{aligned}
\tag{9}
$$

The initial conditions of Eq. (9) are defined for the beginning of a pandemic as

follows:

$$\begin{aligned} &R_\emptyset(0) = N-2,\ R_\emptyset I_1(0) = 1,\ R_\emptyset I_2(0) = 1, \\ &D(0) = R_{\{1\}}(0) = R_{\{2\}}(0) = R_{\{1,2\}}(0) = R_{\{1\}}I_2(0) = R_{\{2\}}I_1(0) = 0 \end{aligned} \tag{10}$$

Moreover,

$$N = R_\emptyset + R_\emptyset I_1 + R_\emptyset I_2 + R_{\{1\}}I_2 + R_{\{2\}}I_1 + R_{\{1\}} + R_{\{2\}} + R_{\{1,2\}} + D. \tag{11}$$

We use a model that does not allow temporary cross-immunity and without increased susceptibility to the second infection.

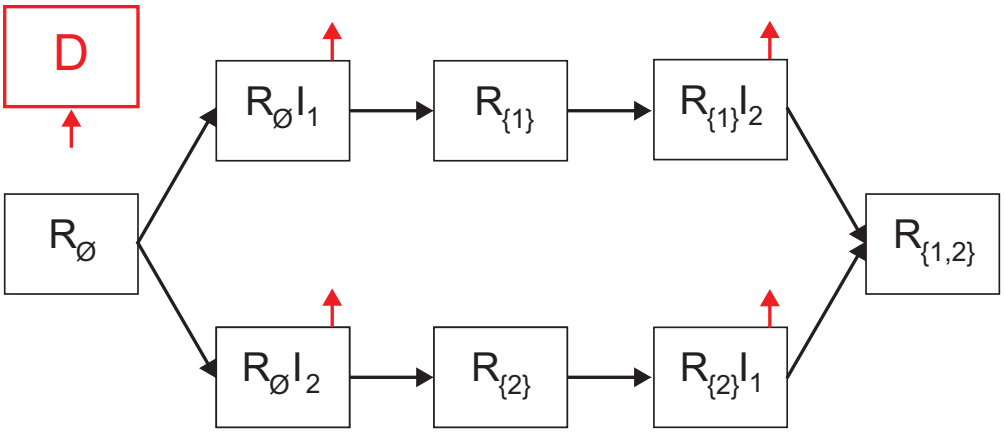

Fig 5. Schematic view of transition between disease stages in the case where $|M| = 2$.

For $|M| = 2$, we are interested in the equilibrium states of the model, especially stable states in which a pandemic can persist for a long time. In addition, we investigate the basic reproduction number (R_0), as it is an indicator of a pandemic outbreak ($R_0 > 1$), and is considered the main characteristic of a pandemic.

3.1 Equilibria

The equilibrium state of the model is the state in which the gradient is zero [49]. Therefore, Eq. (12) takes the form:

$$\begin{aligned}
&-R_\emptyset\big(\beta_{\emptyset,1}(R_\emptyset I_1 + R_{\{2\}}I_1) + \beta_{\emptyset,2}(R_\emptyset I_2 + R_{\{1\}}I_2)\big) = 0,\\
&\beta_{\emptyset,1}(R_\emptyset I_1 + R_{\{2\}}I_1)R_\emptyset - \gamma_{\emptyset,1}R_\emptyset I_1 = 0,\\
&\beta_{\emptyset,2}(R_\emptyset I_2 + R_{\{1\}}I_2)R_\emptyset - \gamma_{\emptyset,2}R_\emptyset I_2 = 0,\\
&\gamma_{\emptyset,1}\phi_{\emptyset,1}R_\emptyset I_1 - \beta_{\{1\},2}(R_{\{1\}}I_2 + R_\emptyset I_2)R_{\{1\}} = 0,\\
&\gamma_{\emptyset,2}\phi_{\emptyset,2}R_\emptyset I_2 - \beta_{\{2\},1}(R_{\{2\}}I_1 + R_\emptyset I_1)R_{\{2\}} = 0,\\
&\beta_{\{1\},2}(R_{\{1\}}I_2 + R_\emptyset I_2)R_{\{1\}} - \gamma_{\{1\},2}R_{\{1\}}I_2 = 0,\\
&\beta_{\{2\},1}(R_{\{2\}}I_1 + R_\emptyset I_1)R_{\{2\}} - \gamma_{\{2\},1}R_{\{2\}}I_1 = 0,\\
&\gamma_{\{2\},1}\phi_{\{2\},1}R_{\{2\}}I_1 + \gamma_{\{1\},2}\phi_{\{1\},2}R_{\{1\}}I_2 = 0,\\
&\gamma_{\emptyset,1}(1-\phi_{\emptyset,1})R_\emptyset I_1 + \gamma_{\{2\},1}(1-\phi_{\{2\},1})R_{\{2\}}I_1\\
&+\gamma_{\emptyset,2}(1-\phi_{\emptyset,2})R_\emptyset I_2 + \gamma_{\{1\},2}(1-\phi_{\{1\},2})R_{\{1\}}I_2 = 0.
\end{aligned} \tag{12}$$

From Eq. (12), the *pandemic-free* equilibria is obtained where

$$R_\emptyset I_1^* = R_\emptyset I_2^* = R_{\{1\}}I_2^* = R_{\{2\}}I_1^* = 0,$$

because in this state, there are no more infected individuals which means all strains have gone extinct. Therefore, the equilibria states take the form:

$$R_\emptyset^* = \mu_1,\ R_{\{1\}}^* = \mu_2,\ R_{\{2\}}^* = \mu_3,\ R_{1,2}^* = \mu_4,\ D^* = N - \Sigma_{i=1}^4\mu_i. \tag{13}$$

According to [49], this set of states (Eq. (13)) is the only asymptotically stable equilibria of the model. Nonetheless, the equilibria states where strain $i = 1$ is over are obtained where

$$R_\emptyset I_1^* = R_{\{2\}}I_1^* = 0.$$

These equilibria have an epidemiological interest as the extinction of one of two strains can be a turning point in pandemic management policies. Thus, is assumes without loss of generality that $i = 1$. Hence, from the fourth and sixth equations, one obtains that

$$RI_2^* = R_{\{1\}}I_2^* = 0.$$

Accordingly, the system converges to the *pandemic-free* equilibria.

In addition, while other equilibria states theoretically exist (by relaxing the previous assumptions), from an epidemiological point of view, the unstable equilibria are obtained in the middle of the pandemic. It is possible to see that it is enough that an individual may recover in order to diverge from each one of these equilibria states. As such, these equilibria obtained, if any, do not provide a meaningful point in the pandemic's dynamics.

3.2 Basic reproduction number

The basic reproduction number, R_0, is defined as the expected number of secondary cases produced by a single (typical) infection in a completely susceptible population [50]. In the case of a *SIR*-based model, the basic reproduction number indicates an epidemic outbreak if $R_0 > 1$ or not if $R_0 < 1$.

To find the basic reproduction number for two strains, we use the Next Generation Matrix (NGM) approach [51]. First, we compute the new infections matrix

$$\mathbf{F} = \begin{pmatrix} \beta_{\emptyset,1}R_{\emptyset} & 0 & \beta_{\{2\},1}R_{\emptyset} & 0 \\ 0 & \beta_{\emptyset,2}R_{\emptyset} & 0 & \beta_{\{1\},2}R_{\emptyset} \\ \beta_{\{2\},1}R_{\{2\}} & 0 & \beta_{\{2\},1}R_{\{2\}} & 0 \\ 0 & \beta_{\{1\},2}R_{\{1\}} & 0 & \beta_{\{1\},2}R_{\{1\}} \end{pmatrix}. \tag{14}$$

Afterward, we compute the transfers of infections from one compartment to another matrix

$$\mathbf{V} = \begin{pmatrix} \gamma_{\emptyset,1} & 0 & 0 & 0 \\ 0 & \gamma_{\emptyset,2} & 0 & 0 \\ 0 & 0 & \gamma_{\{1\},2} & 0 \\ 0 & 0 & 0 & \gamma_{\{2\},1} \end{pmatrix} \rightarrow \mathbf{V}^{-1} = \begin{pmatrix} 1/\gamma_{\emptyset,1} & 0 & 0 & 0 \\ 0 & 1/\gamma_{\emptyset,2} & 0 & 0 \\ 0 & 0 & 1/\gamma_{\{1\},2} & 0 \\ 0 & 0 & 0 & 1/\gamma_{\{2\},1} \end{pmatrix}. \tag{15}$$

Now, R_0 is the dominant eigenvalue of the matrix [51].

$$\mathbf{G} = \mathbf{F}\mathbf{V}^{-1} = \begin{pmatrix} \frac{\beta_{\emptyset,1}R_{\emptyset}}{\gamma_{\emptyset,1}} & 0 & \frac{\beta_{\{2\},1}R_{\emptyset}}{\gamma_{\{1\},2}} & 0 \\ 0 & \frac{\beta_{\emptyset,2}R_{\emptyset}}{\gamma_{\emptyset,2}} & 0 & \frac{\beta_{\{1\},2}R_{\emptyset}}{\gamma_{\{2\},1}} \\ \frac{\beta_{\{2\},1}R_{\{2\}}}{\gamma_{\emptyset,1}} & 0 & \frac{\beta_{\{2\},1}R_{\{2\}}}{\gamma_{\{1\},2}} & 0 \\ 0 & \frac{\beta_{\{1\},2}R_{\{1\}}}{\gamma_{\emptyset,2}} & 0 & \frac{\beta_{\{1\},2}R_{\{1\}}}{\gamma_{\{2\},1}} \end{pmatrix} \tag{16}$$

which is obtained from the root of the representative polynomial:

$$\begin{aligned} 0 = {} & \lambda^4 - \lambda^3\left(\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} + \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}} + \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\right) + \\ & \lambda^2\left(2\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} - \frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}} + \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}} - \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{1\},2}}{\gamma_{\emptyset,2}} + \frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}}\frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}} + \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}} - \frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}}\right) + \\ & \lambda\left(-\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}^2\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} + \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{1\},2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} - 2\frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} + \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}} - \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}} + \right. \\ & \left. \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{1\},2}}{\gamma_{\emptyset,2}} + \frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} + \frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}}\right) + \\ & \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}^2\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} - \frac{\beta_{\emptyset,1}}{\gamma_{\emptyset,1}}\frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{1\},2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}} - \frac{\beta_{\emptyset,2}}{\gamma_{\emptyset,2}}\frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}} + \frac{\beta_{\{2\},1}}{\gamma_{\emptyset,1}}\frac{\beta_{\{2\},1}}{\gamma_{\{1\},2}}\frac{\beta_{\{1\},2}}{\gamma_{\{2\},1}}\frac{\beta_{\{1\},2}}{\gamma_{\emptyset,2}}. \end{aligned} \tag{17}$$

Using the *Matlab*'s (version 2021b) symbolic programming, one is able to obtain the R_0. Just find the roots of the polynomial shown in Eq. (17) and take the biggest one.

This approach cannot be generalized for more than two strains $|M| > 2$ as the NGM will be of size $k \times k$ where $k = \Sigma_{i=1}^{|M|}\binom{n}{x}$. Namely, the size of the NGM is larger than four and according to Galois theory [52] and based on Abel–Ruffini theorem [53], the roots of the representative polynomial of the NGM cannot be obtained using radicals. This means one cannot analytically find the eigenvalues of the NGM which are used to obtain R_0.

3.3 Model validation

The model validation is divided into two phases: parameter estimation and historical fitting. The parameter estimation method allows to use of the proposed model on a specific pandemic and the historical fitting shows the ability of the proposed model to approximate real pandemic spread dynamics given the obtained parameters.

3.3.1 Parameter estimation

The proposed epidemiological model parameter for the case $|M| = 2$ is obtained by fitting the proposed model onto the historical data from April 1 (2020) to December 1 (2020) of UK by the WHO [7], using the fourth-order Runge-Kutta [47] and gradient descent [54] algorithms. These dates are picked as the population in the UK during this period had not been vaccinated against the COVID-19 disease yet and a second strain (i.e., the COVID-19 UK Variant - B.1.1.7) appeared according to [55], which based their analysis on clinical testing and later reverse engineering of the mutation appearance [56]. Both point to the same period while do not have a full agreement on the specific dates of the appearance of the mutation. Specifically, we randomly guess the values of the model's parameters, solving the system of ODEs using the fourth-order Runge-Kutta method and computing the Gaussian (L_2) distance from the historical data. In particular, we used the daily number of infection, recovered, and deceased individuals. As such, the fitness function takes the form:

$$F(H,P)_{[t_0,t_f]} := \sqrt{\Sigma_{t=t_0}^{t_f}\big((H[S](t)-P[S](t))^2+(H[R](t)-P[R](t))^2+(H[I](t)-P[I](t))^2\big)}, \tag{18}$$

where $H[X](t)$ is the historical size of the population at the epidemiological state X at time t and $P[X](t)$ is the model's prediction size of the population at the epidemiological state X at time t. The model's $P[I]$ and $P[R]$ refer to all states for the form R_jI_i and R_j, respectively.

Afterward, we repeated this process while modifying the value of a single parameter by some pre-defined value $\delta = 0.01$, obtaining a numerical gradient. At this stage, we used the gradient descent algorithm in order to find the values that minimize the model's L_2 distance from the historical data using Eq. (18). The process is stopped once the gradient's (L_1) norm is smaller than some pre-defined threshold value $\epsilon = 0.1$. The entire process is repeated $r = 100$ times and the parameter values that are obtained most often are decided to be the model's parameter value. The values for (δ, ϵ, r) are manually picked. A schematic view of the fitting method is presented in Fig. 6.

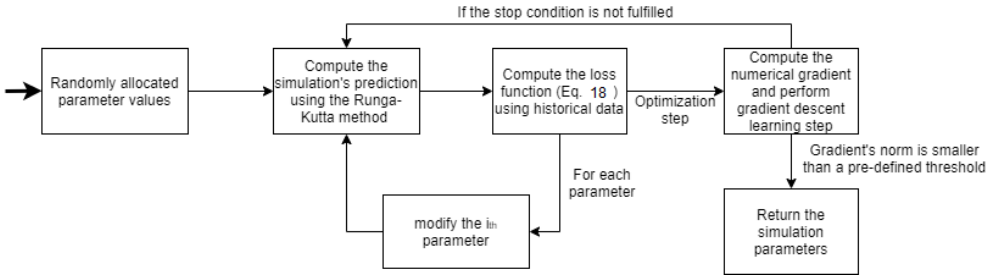

Fig 6. A schematic view of the fitting method.

3.3.2 Historical fitting

In order to numerically evaluate the ability of the proposed model to fit real epidemiological data, we decided to simulate the COVID-19 pandemic in the United Kingdom (UK). This case is chosen due to the availability of epidemiological data and since a COVID-19 strain is known to originate in the UK [7, 57]. Therefore, we computed the parameter values, assuming the initial conditions taking the form:

$$\begin{aligned} &R_\emptyset(0) = 67200000,\ R_\emptyset I_1(0) = 100,\ R_\emptyset I_2(0) = 1,\ D(0) = 0 \\ &R_{\{1\}}(0) = \ R_{\{2\}}(0) = R_{\{1,2\}}(0) = R_{\{1\}}I_2(0) = R_{\{2\}}I_1(0) = 0. \end{aligned} \tag{19}$$

where $R_{\emptyset}(0) = 67200000$ to represent the size of the UK population in the beginning of the pandemic. A summary of the obtained parameter values is shown in Table 1, such that 27% of the random parameter value initial conditions converged to the values with $d_{L_2} = 0.089$. Namely, the model has a daily mean square error of 8.9%.

Parameter Definition	Symbol	Value
Infection rate of the strain ($i = 1$) [1]	$\beta_{\emptyset,1}$	0.04
Infection rate of the strain ($i = 2$) [1]	$\beta_{\emptyset,2}$	0.07
Infection rate of the strain ($i = 2$), after recovery from the strain ($i = 1$) [1]	$\beta_{\{1\},2}$	0.01
Infection rate of the strain ($i = 1$), after recovery from the strain ($i = 2$) [1]	$\beta_{\{2\},1}$	0.02
The average duration that is takes for an individual to recover from the strain ($i = 1$) in days [t^{-1}]	$\gamma_{\emptyset,1}$	0.08
The average duration that is takes for an individual to recover from the strain ($i = 2$) in days [t^{-1}]	$\gamma_{\emptyset,2}$	0.06
The average duration that is takes for an individual to recover from the strain ($i = 1$) after recovering from the strain ($i = 2$) in days [t^{-1}]	$\gamma_{\{2\},1}$	0.21
The average duration that is takes for an individual to recover from the strain ($i = 2$) after recovering from the strain ($i = 1$) in days [t^{-1}]	$\gamma_{\{1\},2}$	0.17
The probability an infected individual will recover from the strain ($i = 1$) [1]	$\phi_{\emptyset,1}$	0.98
The probability an infected individual will recover from the strain ($i = 2$) [1]	$\phi_{\emptyset,2}$	0.96
The probability an infected individual will recover from the strain ($i = 2$) after recovering from the strain ($i = 2$) [1]	$\phi_{\{1\},2}$	0.99
The probability an infected individual will recover from the strain ($i = 3$) after recovering from the strain ($i = 1$) [1]	$\phi_{\{2\},1}$	0.99

Table 1. A summary of the model parameters and values for the case of $|M| = 2$, obtained from the fitting process to the historical WHO COVID-19 data from April 1 (2020) to December 1 (2020).

A fitting dynamics between the historical data (circle, black) and the model's prediction (axes, blue) is shown in Fig. 7, where the x-axis describes the time from September 1 (2020) to December 1 (2020), and the y-axis describes the daily basic reproduction number (R_0). The historical basic reproduction number (R_0) from WHO is computed using the following formula $R_0(t) := \frac{I(t+1)-I(t)}{R(t+1)-R(t)}$.

4 Discussion

We have developed a mathematical model and a computer simulation aiming at establishing the connections between the number of pandemic disease strains and the pandemic's spread in the population for any pathogen, under the epidemiological *SIRD* model. Unlike the previous modeling approaches [12, 38, 40], we have extended the strain diversity for any arbitrary number (m) and did not introduce any pathogen-specific attributes, keeping the model as generic as possible.

We have shown that for the case of only two strains (e.g., $|M| = 2$), the only stable equilibria states are when the pandemic is over for both strains ($R_{\emptyset}I_1^* = R_{\emptyset}I_2^* = R_{\{1\}}I_2^* = R_{\{2\}}I_1^* = 0$), as shown in Section 3.3.2. The result of the

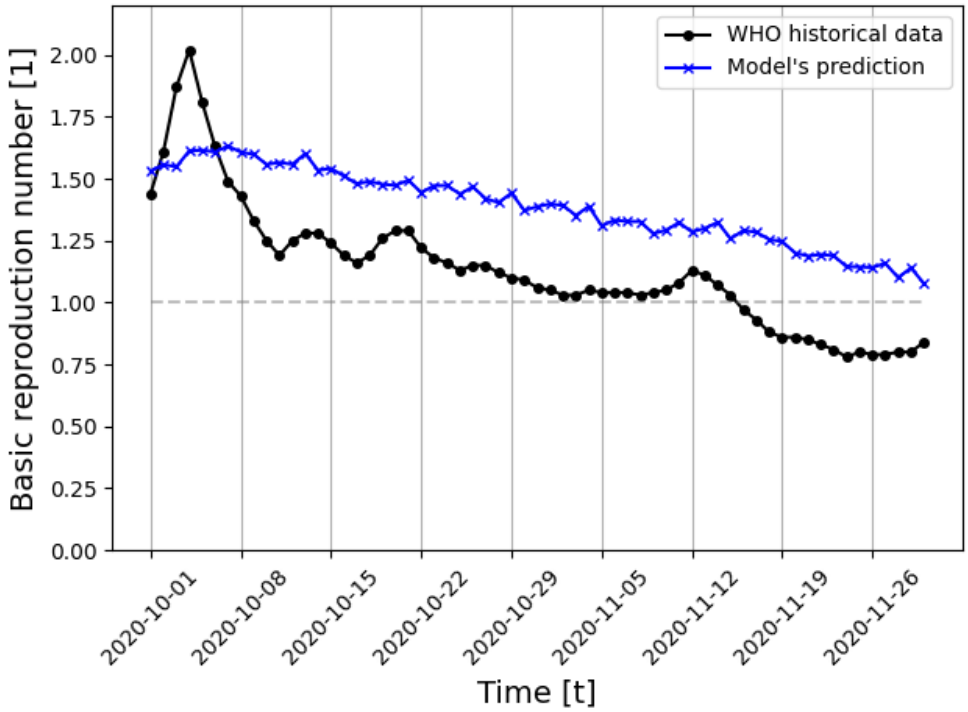

Fig 7. Daily R_0 in UK between September 1 and December 1 (2020) comparison between the historical data (specifically, the daily number of infected, recovered, and dead individuals) and the proposed model predictions (for $|M| = 2$). The gray horizontal line indicates $R_0 = 1$. The model's parameter values are shown in Table 1.

equilibrium analysis is that the *pandemic-free* states are stable only when the epidemics of the two strains cease; that is, after the end of the general pandemic (Eq. (13)).

Moreover, an analytical computation of the basic reproduction number (R_0) requires information on infections between individuals with different strains, which is not realistically available. Therefore, an immediate result of the model is that once a pandemic developed secondary strains, a numerical and statistical approximation of R_0 is left to be the only feasible approach.

In addition, the proposed model is evaluated on the COVID-19 pandemic (for the case of the UK) and has shown promising ability to fit a long period of historical data with multi-strain (eight months, 8.9% daily mean square error). A prediction of the last two months of this period is shown in Fig. 7, based on the obtained model's parameter values which are presented in Table 1. Strain $i = 1$ is mapped to the original strain of COVID-19. Since at the beginning of the pandemic, this was only a single strain, the measured epidemiological values are necessarily associated with this strain. This is not the case for measurements of periods where two or more strains existed. The proposed model captures a general trend of decreasing R_0 during this period while not matching the data closely as it does not take into consideration other social and epidemiological dynamics intentionally, allowing analytical analysis to be considered. However, future extensions of the proposed model should be able to predict more closely historical pandemic events.

According to Voinsky et al. [58], the average recovery rate of strain ($i = 1$) is 0.0714 while the model predicted $\gamma_{\emptyset,1} = 0.08$ (where the approximation size is $\delta = 0.01$), as presented in Table 1. In addition, according to WHO [7], the average mortality rate of

this period is ~ 0.0138 while the model predicted that the average mortality rate from this strain is $1 - \phi_{\emptyset,1} = 0.02$. Thus, while the model is simple, it is able to capture the biological and epidemiological properties of the pandemic.

Furthermore, we evaluated the influence of the number of strains on the mean basic reproduction number ($E[R_0]$), mortality rate, and a maximum number of infected individuals, as shown in Figs. 2, 4, and 3, respectively. We show that the basic reproduction number is upper bounded by taking into consideration only the most aggressive strain. Formally, we perform a paired two-tail T-test in order to evaluate if the processes differ in a statistically significant way with $\alpha = 0.05$ and obtain that 0 is not in the confidence interval of the statistical test. As such, the most aggressive strain approximation to the pandemic can be used as an upper boundary for the mean basic reproduction number. An immediate outcome is that the proposed model is upper bounded by the *SIRD* model with the slight modification that each individual can be infected up to $|M|$ times. This means one can get a statistically similar result (on average) to a pandemic with $|M|$ strains by using a simpler model that requires less biological and epidemiological data compared to the proposed model. These results agree with the analysis performed by Dang et al. [37] on a multi-strain model for influenza.

Based on Eq. (7), the maximum number of infected individuals is growing in a logarithmic manner to the number of strains when the latter occurs simultaneously. In a similar manner, based on Eq. (8), the mortality rate is growing in a logarithmic manner to the number of strains when the latter occur simultaneously. We numerically show in Figs. 3 and 4 that the epidemiological properties which indicate the severity of the pandemic in a well-mixed population grow in a logarithmic manner as a function of the number of strains ($|M|$). This connection indicates that the first few strains make a relatively large contribution to the mortality and pandemic spread dynamics, but as the number of strains grows, each strain contributes less to these numbers. Policymakers can take advantage of this link when planning intervention policies to contain the spread of a pandemic, given that new strains can emerge during pathogen mutation. The code developed for this model is publicly available as an open source[2].

Several possible future research directions emerge from the proposed initial modeling. First, one can introduce a fixed delay parameter to the occurrence of strains, investigating the influence of this parameter on the epidemiological spread. Second, one can take into consideration more detailed biological settings, assuming the stochastic occurrence of the strains from some distribution. Both directions aim to better represent a real pandemic where several strains do not exist from the beginning of the pandemic. Moreover, one can introduce a similarity matrix between the strains as they are mutations of an original strain, which are reflected by the immunity response to re-infection of different strains or a cross-immunity response. In the same direction, adding an Exposed state would make the proposed model more biologically accurate, since most strains have an incubation period before the host becomes infectious. The multi-strain model is a theoretical platform that will help guide the decision-making process in the event of a pandemic crisis while providing the forecast of the results of the selected course of action.

References

1. Janku A, Schenk G, Mauelshagen F. Historical Disasters in Context: Science, Religion, and Politics. Routledge; 2011.

[2]ADD HERE LATER

2. Gottshang TR. Economic Change, Disasters, and Migration: The Historical Case of Manchuria. Economic Development and Cultural Change. 1987;45(3).

3. Noji EK, Toole MJ. The Historical Development of Public Health Responses to Disasters. Disasters. 1997;21:366–376.

4. van Bavel BJP, Curtis DR. Better Understanding Disasters by Better Using History. International Journal of Mass Emergencies and Disasters. 2016;34(1):143–169.

5. Conti AA. Historical and methodological highlights of quarantine measures: from ancient plague epidemics to current coronavirus disease (COVID-19) pandemic. Acta bio-medica : Atenei Parmensis. 2020;91(2):226–229.

6. Brodeur A, Gray D, Islam A, Bhuiyan S. A Literature Review of the Economics of COVID-19. IZA Discussion Paper No 13411, Available at SSRN: https://ssrncom/abstract=3636640. 2020;.

7. WHO. WHO Coronavirus Disease (COVID-19) Dashboard;. Available from: `https://covid19.who.int/`.

8. Lazebnik T, Bunimovich-Mendrazitsky S, Shami L. Pandemic management by a spatio–temporal mathematical model. International Journal of Nonlinear Sciences and Numerical Simulation. 2021;.

9. Lederberg J. Medical Science, Infectious Disease, and the Unity of Humankind. JAMA. 1988;260(5):684–685.

10. Wu T, Perrings C, Kinzig A, Collins JP, Minteer BA, Daszak P. Economic growth, urbanization, globalization, and the risks of emerging infectious diseases in China: A review. Ambio. 2017;46(1):18–29.

11. Cheng XX, Wang Y, Huang G. Dynamics of a competing two-strain SIS epidemic model with general infection force on complex networks. Nonlinear Analysis-Real World Applications. 2021;59:103247.

12. Gordo I, Gomes MGM, Reis DG, Campos PRA. Genetic Diversity in the SIR Model of Pathogen Evolution. Plos One. 2009;4(3):e4876.

13. Shi P, Keskinocak P, Swann J, Lee B. Modelling seasonality and viral mutation to predict the course of an influenza pandemic. Epidemiology and Infection. 2010;138(10):1472–1481.

14. Kermack WO, McKendrick AG. A contribution to the mathematical theory of epidemics. Proceedings of the Royal Society. 1927;115:700–721.

15. Libi F, Weiguo S, Wei L, Siuming L. Simulation of emotional contagion using modified SIR model: A cellular automaton approach. Physica A: Statistical Mechanics and its Applications. 2014;405(1):380–391.

16. Lazebnik T, Shami L, Bunimovich-Mendrazitsky S. Spatio-Temporal Influence of Non-Pharmaceutical Interventions Policies on Pandemic Dynamics and the Economy: The Case of COVID-19. Epidemiologic-Economic. 2021;.

17. Bognanni M, Hanley D, Kolliner D, Mitman K. Economics and Epidemics: Evidence from an Estimated Spatial Econ-SIR Model PDF Logo. Institute of Labor Economics. 2020;.

18. Milner FA, Zhao R. S-I-R Model with Directed Spatial Diffusion. An International Journal of Mathematical Demography. 2008;15(3).

19. Oka T, Wei W, Zhu D. A Spatial Stochastic SIR Model for Transmission Networks with Application to COVID-19 Epidemic in China. SSRN. 2020;.

20. Lazebnik T, Alexi A. Comparison of Pandemic Intervention Policies in Several Building Types Using Heterogeneous Population Model. medrxiv. 2021;.

21. Towers S, Vogt Geisse K, Zheng Y, Feng Z. Antiviral treatment for pandemic influenza: Assessing potential repercussions using a seasonally forced SIR model. Journal of Theoretical Biology. 2011;289:259–268.

22. Kamp C, Heiden M, Henseler O, Seitz R. Management of blood supplies during an influenza pandemic. Transfusion. 2010;50:231–239.

23. Weiss H. The SIR model and the Foundations of Public Health. Materials Matemàtics. 2013; p. 1–17.

24. Cooper I, Mondal A, Antonopoulos CG. A SIR model assumption for the spread of COVID-19 in different communities. Chaos, Solitons Fractals. 2020;139:110057.

25. Agarwal M, Bhadauria AS. Modeling Spread of Polio with the Role of Vaccination. Applications and Applied Mathematics. 2011;6:552 – 571.

26. Shi P, Keskinocak P, Swann JL, Lee BY. Modelling seasonality and viral mutation to predict the course of an influenza pandemic. Epidemiology and Infection. 2010;138(10):1472–1481.

27. Bunimovich-Mendrazitsky S, Stone L. Modeling polio as a disease of development. Journal of Theoretical Biology. 2005;237:302–315.

28. Wells VR, Plotch SJ, DeStefano JJ. Determination of the mutation rate of poliovirus RNA-dependent RNA polymerase. Virus Research. 2001;74(1):119–132.

29. Reluga TC. An SIS epidemiology game with two subpopulations. Journal of Biological Dynamics. 2008; p. 515–531.

30. Yicang Z, Hanwu L. Stability of periodic solutions for an SIS model with pulse vaccination. Mathematical and Computer Modelling. 2003;38(3):299–308.

31. Yanli Z, Sanling Y, Dianli Z. Threshold behavior of a stochastic SIS model with Levy jumps. Applied Mathematics and Computation. 2016;275:255–267.

32. Rappuoli R, Dormitzer PR. Influenza: Options to Improve Pandemic Preparation. Science. 2012;336(6088):1531–1533.

33. Tkachenko AV, Maskov S, Elbanna A, Wong GN, Weiner ZJ, Goldenfeld N. Time-dependent heterogeneity leads to transient suppression of the COVID-19 epidemic, not herd immunity. Science. 2021;118(17):e2015972118.

34. Moein S, Nickaeen N, Roointan A, Bohani N, Heidary Z, Javanmard SH, et al. Inefficiency of SIR models in forecasting COVID-19 epidemic: a case study of Isfahan. Scientific Reports. 2021;11:4725.

35. Chowell G, Sattenspiel L, Bansal S, Viboud C. Mathematical models to characterize early epidemic growth: A review. Physics of Life Reviews. 2016;18:66–97.

36. Minayev P, Ferguson N. Improving the realism of deterministic multi-strain models: implications for modelling influenza A. Journal of the Royal Society Interface. 2008;.

37. Dang YX, Li XZ, Martcheva M. Competitive exclusion in a multi-strain immuno-epidemiological influenza model with environmental transmission. Journal of Biological Dynamics. 2016;10(1).

38. Marquioni VM, de Aguiar MAM. Modeling neutral viral mutations in the spread of SARS-CoV-2 epidemics. Plos One. 2021;16(7):e0255438.

39. Khyar O, Allali K. Global dynamics of a multi-strain SEIR epidemic model with general incidence rates: application to COVID-19 pandemic. Nonlinear Dynamics. 2020;102:489–509.

40. Gubar E, Taynitskiy V, Zhu Q. Optimal Control of Heterogeneous Mutating Viruses. Games. 2018;9(4):103.

41. Aleta A, Hisi ANH, Meloni S, Poletto C, Colizza V, Moreno Y. Human mobility networks and persistence of rapidly mutating pathogens. Royal Society Open Science. 2017;4:160914.

42. Di Giamberardino P, Iacoviello D, Papa F, Sinisgalli C. A data-driven model of the COVID-19 spread among interconnected populations: epidemiological and mobility aspects following the lockdown in Italy. Nonlinear Dynamics. 2021;106:1239–1266.

43. Cox RJ, Brokstad KA. Not just antibodies: B cells and T cells mediate immunity to COVID-19. Nature Reviews Immunology. 2020;20:581–582.

44. Winer DA, Winder S, Shen L, Wadia PP, Yantha J, Paltser G, et al. B-cells promote insulin resistance through modulation of T cells and production of pathogenic IgG antibodies. Nature Medicine. 2011;17:610–617.

45. Mukherjee S, Tworowski D, Detroja R, Mukherjee SB, Frenkel-Morgenstern M. Immunoinformatics and structural analysis for identification of immunodominant epitopes in SARS-CoV-2 as potential vaccine targets. Vaccines. 2020;.

46. Yaqinuddin A. Cross-immunity between respiratory coronaviruses may limit COVID-19 fatalities. Medical hypotheses. 2020;.

47. Yaakub AR, Evans DJ. A fourth order Runge–Kutta RK(4,4) method with error control. International Journal of Computer Mathematics. 1996; p. 383–411.

48. Bjorck A. Numerical Methods for Least Squares Problems. Society for Industrial and Applied Mathematics. 1996;5:497–513.

49. Lazebnik T, Bunimovich-Mendrazitsky S, Shaikhet L. Novel Method to Analytically Obtain the Asymptotic Stable Equilibria States of Extended SIR-type Epidemiological Models. Symmetry. 2021;.

50. van den Driessche P, Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Mathematical Biosciences. 2002;180(1):29–48.

51. Diekmann O, Heesterbeek JA, Roberts MG. The construction of next-generation matrices for compartmental epidemic models. Journal of the Royal Society. 2010;7(47):873–885.

52. Galois E, Neumann PM. The Mathematical Writings of Évariste Galois. European Mathematical Society. 2011;.

53. Abel NH. Mémoire sur les équations algébriques, ou l'on démontre l'impossibilité de la résolution de l'équation générale du cinquième degré. Sylow, Ludwig; Lie, Sophus. 1824; p. 28–33.

54. Curry HB. The method of steepest descent for non-linear minimization problems. Quarterly of Applied Mathematics. 1944;2(3):258–261.

55. Mahase E. Covid-19: What have we learnt about the new variant in the UK? BMJ. 2020;.

56. Sherman SM, Smith LE, Sim J, Amlot R, Cutts M, Dasch H, et al. COVID-19 vaccination intention in the UK: results from the COVID-19 vaccination acceptability study (CoVAccS), a nationally representative cross-sectional survey. Human Vaccines Immunotherapeutics. 2021;6:1612–1621.

57. Wise J. Covid-19: New coronavirus variant is identified in UK. BMJ. 2020;371.

58. Voinsky I, Baristaite G, Gurwitz D. Effects of age and sex on recovery from COVID-19: Analysis of 5769 Israeli patients. The Journal of infection. 2020;81:102–103.