

## RESEARCH ARTICLE

# Analysis of multi-strain infection of vaccinated and recovered population through epidemic model: Application to COVID-19

Olusegun Michael Otunuga \*

Department of Mathematics, Augusta University, Augusta, GA, United States of America

\* [ootunuga@augusta.edu](mailto:ootunuga@augusta.edu)

## Abstract

In this work, an innovative multi-strain *SEAIR* epidemic model is developed for the study of the spread of a multi-strain infectious disease in a population infected by mutations of the disease. The population is assumed to be completely susceptible to  $n$  different variants of the disease, and those who are vaccinated and recovered from a specific strain  $k$  ( $k \leq n$ ) are immune to previous and present strains  $j = 1, 2, \dots, k$ , but can still be infected by newer emerging strains  $j = k + 1, k + 2, \dots, n$ . The model is designed to simulate the emergence and dissemination of viral strains. All the equilibrium points of the system are calculated and the conditions for existence and global stability of these points are investigated and used to answer the question as to whether it is possible for the population to have an endemic with more than one strain. An interesting result that shows that a strain with a reproduction number greater than one can still die out on the long run if a newer emerging strain has a greater reproduction number is verified numerically. The effect of vaccines on the population is also analyzed and a bound for the herd immunity threshold is calculated. The validity of the work done is verified through numerical simulations by applying the proposed model and strategy to analyze the multi-strains of the COVID-19 virus, in particular, the Delta and the Omicron variants, in the United State.

## OPEN ACCESS

**Citation:** Otunuga OM (2022) Analysis of multi-strain infection of vaccinated and recovered population through epidemic model: Application to COVID-19. PLoS ONE 17(7): e0271446. <https://doi.org/10.1371/journal.pone.0271446>

**Editor:** Mohammed S. Abdo, Hodeidah University, YEMEN

**Received:** February 15, 2022

**Accepted:** June 30, 2022

**Published:** July 29, 2022

**Copyright:** © 2022 Olusegun Michael Otunuga.

This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** Weekly US COVID-19 Vaccine and the SARS-CoV-2 weekly Variant Proportions are available freely on the Centers for Disease Control and Prevention (CDC) websites listed below: <https://data.cdc.gov/Vaccinations/> <https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-Jurisdiction/unsk-b7fch> <https://covid.cdc.gov/covid-data-tracker/#monitoring-variant-heading>.

**Funding:** The author(s) received no specific funding for this work.

## 1 Introduction

The growing threat of infectious diseases with resistance to drugs and vaccinations, causing large number of deaths worldwide, is a cause for concern to the medical community and the general population. Scientists around the world are working to learn more about such diseases in order to study how likely an emerging variant of the disease can spread more easily than existing original variants. More data and analyses are needed for such study. These analyses will shed more light on the possibility of reinfections in people who already recovered from original strain, and infections in people who are fully vaccinated against original or previous strains. An example of such disease is an emerging virus called the corona virus 2019 (COVID-19) virus that has infected and killed millions around the world within a period of two years. The virus was caused by the virus species 'severe acute respiratory syndrome corona virus', named SARS-CoV-2. The airborne transmission occurs by inhaling droplets loaded

**Competing interests:** The authors have declared that no competing interests exist.

with SARS-CoV-2 particles that are expelled by infectious people. Symptoms of the virus appear 2–14 days of exposure to certain strains of the virus. Several SARS-CoV-2 variants have emerged around the world, with each new variants having different characteristics. The COVID-19 virus evolves as changes in its genetic code occur during replication of the genome. The United States Centers for Disease Control and Prevention (CDC) and the US government SARS-COV-2 Inter-agency Group (SIG) [1] have confirmed the emergence of at least twelve new variants of the virus. The SIG group evaluates the risk of the circulating SARS-CoV-2 variants in the United States and make recommendations about the impact, severity, and how spread the virus is. Lineages of these variants that are classified as variant being monitored (VBM) and designated as Variant of Concerned (VOC) in the United States may lead to more severe cases of the virus. By lineages, we mean a group of related virus variant from a common ancestor. The variant, called Omicron-B.1.1.529 [2] was first detected in specimens collected on November 11, 2021 in Botswana and on November 14, 2021 in South Africa. It was first detected in the United States on December 1, 2021 and classified as a VOC. There are increased attention given to the Omicron variant. It is now the main circulating variant in the United States as of early January, 2022, accounting for about 95% of the US reported cases because of its high infection rate. Studies show that even vaccinated individuals can still be infected with this new variant if proper care is not taken. Scientist around the world are working to learn more about this new variant, and to gather more data for the purpose of studying if the variant causes more illness, hospitalizations, and death than infection with other variants. As of early December 2021, one of the new variants, called the Delta-B.1.617.2 variant was said to be the main circulating variant in the United States, and also classified as variant of concern by the SIG group. It was first identified in India and was reported to spread much faster and causes more severe cases than other early variants in the United States, probably causing twice [3] as many infections. Other variants, called the Alpha-B.1.1.7, Beta-B.1.351, Gamma-P.1, Epsilon B.1.427 & B.1.429, Eta B.1.525, Iota B.1.526, Kappa B.1.617.1, Zeta P.2, Mu B.1.621, are classified as VBM. The Alpha-B.1.1.7, Beta-B.1.351, and Gamma-P.1 variants were first discovered in the United Kingdom, South Africa, and Japan/Brazil, respectively. We aim to study how these variants are being transmitted and the impact that vaccines are having on mitigating the number of infection cases in a particular population.

Several mathematical models [4–36] have been developed to study the transmission of infectious diseases. Some of these works [5, 11, 20, 24, 31, 32, 37–43] discussed the transmission of infectious disease caused by the variants and lineages of the COVID-19 virus. SIR models including a modified SIR model with two strains and vaccinated group [10], a generalized SIR model with  $n$ - strains [39], a SIR model with complete cross-protection and nonlinear force of infection [44], a coupled multi-strain SIR epidemic model [24] have been developed to describe the transmission of the COVID-19 strains. Other works such as a multi-strain SEIR models with optimal control [5, 6], multi-strain SEIR models with saturated and general incidence rates [5, 15], SIRD [45] and SEIPAHRF model [46] with Caputo fractional derivative, SCIRP model incorporating media influence [47], and statistical analysis [48] have also been considered in describing the transmission of the virus and its strains. We direct the readers to the work of Hattaf et al. [49, 50] for more recent information about the fractional differential equations and its generalization.

Viruses undergo changes and these changes are cause for concern for people who have recovered from the virus, and also to those who are vaccinated against certain strain of the virus. As discussed in Fudolig et al. [10], a highly infectious emergent strain can infect the susceptible population before the original strain, thereby impeding the spread of the original strain or causing the two strains to coexist in an endemic equilibrium. For this reason, the need to determine conditions in which a newly emerged strain and an existing strains that have a means of immunity will coexist in a population is of utmost important.

Most of the papers mentioned above often utilize simpler versions of the multi-strain network, and provides less mathematical details on the asymptotic behavior of the model. In this work, we develop a multi-strain compartmental model by assuming a population is completely susceptible to  $n$ - different variants of a particular virus at the beginning of an epidemic, with the population of the region partitioned into compartments consisting of susceptible population, population vaccinated against strain  $k$  ( $1 \leq k \leq n$ ) of the virus, population exposed to strain  $k$  of the virus, infected asymptomatic population with strain  $k$  of the virus, infected symptomatic population with strain  $k$  of the virus, and population that recovered from strain  $k$  of the infection. By denoting  $P = \{1, 2, \dots, n\}$  and  $\mathcal{S}_r \in 2^P$  as a subset of the power set  $2^P$  with  $r$  number of strains,  $r = 0, 1, \dots, n$ , with  $r = 0$  representing disease-free case, we study the existence and stability conditions for the equilibrium point corresponding to the scenario where only strains in  $\mathcal{S}_r$  survive. This result is used to estimate the secondary number of infections produced by strain- $k$  infected individual when introduced into a completely susceptible population. This number helps public health expert control the spread of the virus as new variants emerge. Conditions under which an endemic with more than one strain of the virus exist are calculated. The question as to whether this is possible is answered by studying the stability analysis of endemic strains. This model also helps us to understand the correlation between the daily number of administered vaccines and the number of infected, exposed, and recovered population in the region better. To understand the disease dynamics better, we study a powerful quantitative concept that can be used to characterize the contagiousness of each strain of the infectious disease and how transmissible they are. The basic reproduction number, which is the expected number of secondary cases produced by a typical infectious individual in a completely susceptible population in the presence and absence of vaccination are calculated. This study also helps to shed more light on the possibility of a disease being eliminated from a population if enough individuals are immune due to either vaccination or recovery from prior exposure to the disease. A bound for the herd immunity threshold, which is the minimum proportion of the population that must be vaccinated in order to stop the spreading of the disease in the population is calculated and analyzed for each variants.

This paper is organized as follows: In Section 2, a model is developed for the transmission of multi strain infectious diseases for the case where individuals vaccinated against specific strains are immune to that strain and its predecessors but can still be infected by newer emerging strains. The validity of the model, together with the existence and uniqueness of its solution is proved in this section. The reproduction numbers for the cases where the population is vaccinated and when it is not vaccinated are calculated. With these, the effect of vaccination in mitigating infection is studied. Existence of equilibrium points for the case where certain strains persist in the population is examined. In Section 3, the local and global stability of all equilibrium points for model (1) is investigated. In addition to the assumption made in Section 2, an extension of model (1) for the case where those that recovered from a particular strain can get infected by emerging strain is derived and analyzed in Section 4. Existence and stability of equilibrium points for the model is also discussed. Numerical simulations are carried out in Section 5 by applying models (1) and (35) to analyze COVID-19 data. Summary of the work done in this work is discussed in Section 6.

## 2 Materials and methods

### 2.1 Model formulation

By assuming a population is completely susceptible to  $n$ - different variants of a particular virus at the beginning of an epidemic, we partitioned the population of the region into compartments consisting of susceptible (denoted  $S(t)$ ) population, population  $V_k(t)$  vaccinated against

**Table 1. Description of variables for the epidemic model.**

Variable	Description
$S$	Population of susceptible individuals
$V_k$	Population of vaccinated individuals immune to strain $k$ of the infection
$E_k$	Population of individuals exposed to strain $k$ of the infection
$A_k$	Population of asymptomatic individuals infected with strain $k$ of the infection
$I_k$	Population of symptomatic individuals infected with strain $k$ of the infection
$R_k$	Population of individuals who recovered from the $k$ -th strain

<https://doi.org/10.1371/journal.pone.0271446.t001>

strain  $k$  of the virus, population  $E_k(t)$  of exposed individuals infected with strain/variant  $k$  of the virus, infected asymptomatic population  $A_k(t)$  with strain/variant  $k$  of the virus, infected symptomatic population  $I_k(t)$  with strain/variant  $k$  of the virus, and population  $R_k(t)$  that recovered from strain  $k$  of the virus at a given time  $t$ , for  $k = 1, 2, \dots, n$ . These state variables are described in Table 1. We assume that the order of existence of newer strains follow  $k = 1, 2, \dots, n$ . That is, the population is first infected by strain 1, followed by strain 2,  $\dots, n$ . We assume those who are vaccinated and recovered from strain  $k$  are immune to previous and present strains  $j = 1, 2, \dots, k$ , but can be infected by newer strains  $j = k + 1, k + 2, \dots, n$ . We can see, for example, that this assumption is satisfied in the case of COVID-19 and supported by the United States Centers for Disease Control and Prevention (CDC), who claimed that current vaccines approved in the United States are expected to protect against severe illness, hospitalizations, and deaths caused by variants of the virus, although people who are fully vaccinated can still be infected. We assume that the migration rate into the susceptible population not receiving vaccination against any of the strains of the virus is  $(1 - q)\mu$ , while the migration rate into the population vaccinated against strain  $k$  is  $\mu q_k$ , where  $q = \sum_{k=1}^n q_k < 1$ . The parameter  $\beta_j$  denotes the per capita contact rate of susceptible individuals with symptomatic infected strain  $j$ ;  $\gamma_j$  denotes the per capita contact rate of susceptible individuals with asymptomatic infected strain  $j$ ;  $\bar{\beta}_{j,k}$  and  $\bar{\gamma}_{j,k}$  denote the reduced per capita contact rates of symptomatic and

**Table 2. Description of parameters for the epidemic model.**

Parameter	Description
$\beta_k$	Transmission rate of symptomatic infected individuals with strain $k$ of infection interacting with susceptible population
$\gamma_k$	Transmission rate of asymptomatic infected individuals with strain $k$ of infection interacting with susceptible population
$\bar{\beta}_{k,j}$	Transmission rate of symptomatic infected individuals with strain $k$ interacting with population vaccinated against strain $j$ of infection, $k = j + 1, \dots, n$
$\bar{\gamma}_{k,j}$	Transmission rate of asymptomatic infected individuals with strain $k$ interacting with population vaccinated against strain $j$ of infection, $k = j + 1, \dots, n$
$h_{k,j}$	Transmission rate of symptomatic infected individuals with strain $k$ interacting with population that recovered from strain $j$ of infection, $k = j + 1, \dots, n$
$\epsilon_{k,j}$	Transmission rate of asymptomatic infected individuals with strain $k$ interacting with population that recovered from strain $j$ of infection, $k = j + 1, \dots, n$
$\mu$	Natural birth rate
$p$	Fraction of infection cases that are asymptomatic
$q_k$	Fraction of population vaccinated against strain $k$ of infection
$\lambda_k$	Transition rate of individuals with strain $k$ of infection from exposed to infected class
$r_k$	Asymptomatic recovery rate of those with strain $k$ of infection
$\theta_k$	Symptomatic recovery rate of those with strain $k$ of infection
$q$	$\sum_{j=1}^n q_j$

<https://doi.org/10.1371/journal.pone.0271446.t002>

asymptomatic strain  $j$  infected individual, respectively, that are vaccinated against strain  $k, j = k + 1, \dots, n; h_{j, k}$  and  $\epsilon_{j, k}$  denote the per capita contact rates of symptomatic and asymptomatic strain  $j$  infected individual, respectively, that recovered from strain  $k, j = k + 1, \dots, n$ . These parameters are described in detail in Table 2.

In this section, we study a case where individuals vaccinated against specific strains are immune to that strain and its predecessors but can still be infected by newer emerging strains. Without loss of generality, we assume the population is normalized so that the sizes  $S, V_k, E_k, A_k, I_k,$  and  $R_k$  are in percentages, for  $k = 1, 2, \dots, n$ . The model governing the transmission of the infectious disease for this case follows the system of deterministic differential equation

$$\begin{aligned}
 dS &= \left( (1 - q)\mu - S \sum_{j=1}^n (\beta_j I_j + \gamma_j A_j) - \mu S \right) dt, \quad S(t_0) = S_0, \\
 dV_k &= \left( q_k \mu - V_k \sum_{j=k+1}^n (\bar{\beta}_{jk} I_j + \bar{\gamma}_{jk} A_j) - \mu V_k \right) dt, \quad V_k(t_0) = V_{k0}, \quad k = 1, 2, \dots, n - 1 \\
 dV_n &= (q_n \mu - \mu V_n) dt, \quad V_n(t_0) = V_{n0}, \\
 dE_1 &= (S(\beta_1 I_1 + \gamma_1 A_1) - e_1 E_1) dt, \quad E_1(t_0) = E_{10}, \\
 dE_k &= \left( S(\beta_k I_k + \gamma_k A_k) + \sum_{j=1}^{k-1} V_j (\bar{\beta}_{kj} I_k + \bar{\gamma}_{kj} A_k) - e_k E_k \right) dt, \quad E_k(t_0) = E_{k0}, \quad k = 2, 3, \dots, n \\
 dA_k &= (p \lambda_k E_k - a_k A_k) dt, \quad k = 1, 2, \dots, n \quad A_k(t_0) = A_{k0}, \quad k = 1, 2, \dots, n \\
 dI_k &= ((1 - p) \lambda_k E_k - w_k I_k) dt, \quad I_k(t_0) = I_{k0}, \quad k = 1, 2, \dots, n, \\
 dR_k &= (\theta_k I_k + r_k A_k - \mu R_k) dt, \quad R_k(t_0) = R_{k0}, \quad k = 1, 2, \dots, n - 1 \\
 dR_n &= (\theta_n I_n + r_n A_n - \mu R_n) dt, \quad R_n(t_0) = R_{n0},
 \end{aligned} \tag{1}$$

where

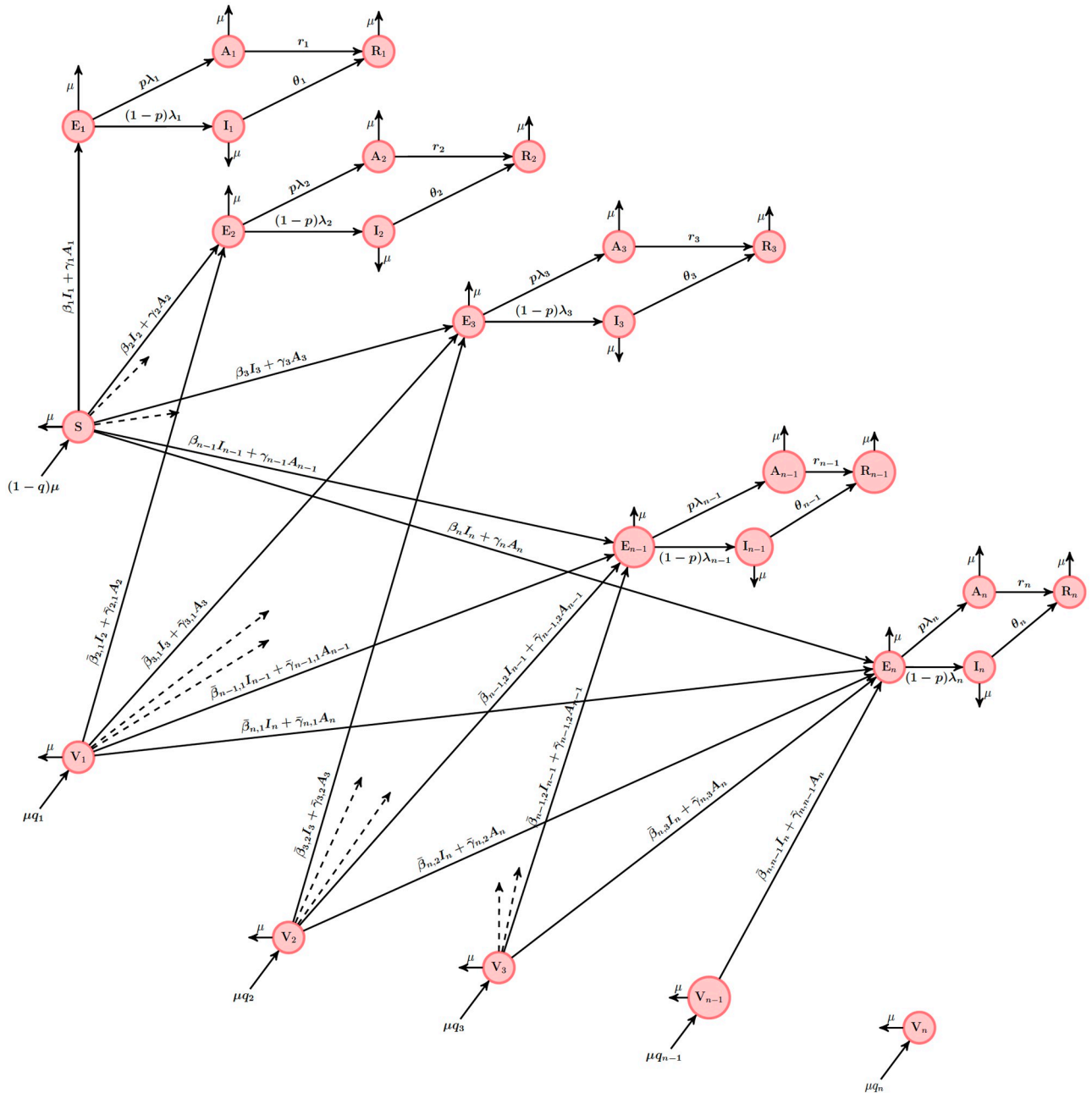
$$\begin{aligned}
 e_k &= \mu + \lambda_k, \\
 a_k &= \mu + r_k, \\
 w_k &= \mu + \theta_k,
 \end{aligned} \tag{2}$$

and the states and parameters in the model are described in Tables 1 and 2, respectively.. An extension of model (1) to include a case where those that recovered from a particular strain can get infected by emerging strain is discussed in Section 4.

In order to better understand the transmission dynamics described in (1), we give a schematic diagram of the model in Fig 1. The circle compartments represent group of individuals. An arrow pointing out of a compartment represents migration of individuals out of the compartment.

**Remark 1.** *Since individuals in the vaccinated group  $V_j$  are only assumed to be immune to strain  $j$  and its predecessors, and not to future strains  $k > j$ , we assume they are also susceptible to future strains  $k > j$ . For this reason, we expect that the rate at which an infectious individual with strain  $k$  make contact with susceptible individuals and individuals in group  $V_j, j < k$ , should be the same. That is,  $\beta_k = \bar{\beta}_{k,j}$  for  $j = 1, 2, \dots, k - 1$ . Likewise, we expect  $\beta_k = h_{k,j}$*

*$\gamma_k = \bar{\gamma}_{k,j} = \epsilon_{k,j}$ , for  $j = 1, 2, \dots, k - 1$ . Vaccination against past and current strains do not*



**Fig 1. Schematic diagram for the epidemic model (1).** The circle compartments represent group of individuals.

<https://doi.org/10.1371/journal.pone.0271446.g001>

provide any protection against new emerging strains. In the case where the infectivity of strain  $k$  ( $k > j$ ) is different for the vaccinated group  $V_j$  due to certain circumstances so that  $\beta_k \neq \bar{\beta}_{k,j}$  and  $\gamma_k \neq \bar{\gamma}_{k,j}$ , we present a comparison of the reproduction number for the case where  $\beta_k = \bar{\beta}_{k,j}$ ,  $\gamma_k = \bar{\gamma}_{k,j}$  and the case where  $\beta_k > \bar{\beta}_{k,j}$ ,  $\gamma_k > \bar{\gamma}_{k,j}$  for  $j = 1, 2, \dots, k - 1$  in Remark 3. Also, based on model (1), an individual infected with certain strain  $k$  cannot be infected with another strain  $j \neq k$  at a given time  $t$ .

### 2.2 Validity of the epidemic model (1)

In this section, we discuss the validity of the proposed epidemic model (1), that is, we discuss the existence and uniqueness of the solution of (1). Let  $N = S + \sum_{k=1}^n (V_k + E_k + A_k + I_k + R_k)$ .

Since the population is assumed to be normalized, we have  $N_0 = N(t_0) = S_0 + \sum_{k=1}^n (V_{k0} + E_{k0} + A_{k0} + I_{k0} + R_{k0}) = 1$  denoting the entire population. It follows that  $dN/dt = \mu - \mu N$ ,  $N(t_0) = 1$ , so that the population size is constant over time. This suggests that the birth and death rates (with the recruitment rate into the susceptible  $S$  and vaccinated  $V_k$  classes in the absence and presence of vaccines being  $(1 - q)\mu$  and  $q_k\mu$ , respectively,  $k = 1, 2, \dots, n$ ) are assumed to be the same so that the population is constant over certain period of time. The most appropriate epidemiological feasible region where solution exist is the set

$$\mathcal{T} = \left\{ (S, \{V_k\}_{k=1}^n, \{E_k\}_{k=1}^n, \{A_k\}_{k=1}^n, \{I_k\}_{k=1}^n, \{R_k\}_{k=1}^n) \in \mathbb{R}_+^{5n+1} \mid 0 \leq S + \sum_{k=1}^n (V_k + E_k + A_k + I_k + R_k) \leq 1 \right\}. \tag{3}$$

The existence, uniqueness, and positiveness of the solution of (1) is shown for the case where  $S_0 > 0, V_{k0} > 0, E_{k0} > 0, A_{k0} > 0, I_{k0} > 0, R_{k0} > 0$  using results from Kelley and Peterson [51].

**Theorem 1.** *If  $S_0 > 0, V_{k0} > 0, E_{k0} > 0, A_{k0} > 0, I_{k0} > 0, R_{k0} > 0, k = 1, 2, \dots, n$ , then there exist a positive unique solution of (1) in the feasible region  $\mathcal{T}$  for all  $t \geq 0$ .*

*Proof.* Let  $S_0 > 0, V_{k0} > 0, E_{k0} > 0, A_{k0} > 0, I_{k0} > 0, R_{k0} > 0$  for model (1). Define

$$\tilde{y} = (S, V_1, \dots, V_n, E_1, \dots, E_n, A_1, \dots, A_n, I_1, \dots, I_n, R_1, \dots, R_n)^T$$

so that Eq (1) can be written as

$$d\tilde{y} = h(t) dt, \quad \tilde{y}(t_0) = \tilde{y}_0, \tag{4}$$

for some vector function  $h(t)$ . It can be shown that  $h : \mathbb{R}^{5n+1} \rightarrow \mathbb{R}^{5n+1}$  is continuous with continuous first-order partial derivatives with respect to  $\tilde{y}_1, \tilde{y}_2, \dots, \tilde{y}_{5n+1}$ . It follows from Theorem 3.1 of Kelley and Peterson [51] that there exist a unique solution  $S(t), V_k(t), E_k(t), A_k(t),$

$I_k(t), R_k(t), k = 1, 2, \dots, n$  of (1) for all  $t \geq 0$ . Let  $\lambda(t) = \sum_{j=1}^n (\beta_j I_j(t) + \gamma_j A_j(t))$ . It follows from

(1) that  $S(t)$  satisfies  $S(t) = S_0 \exp(-\int_0^t (\mu + \lambda(v)) dv) + (1 - q)\mu \exp(-\int_0^t (\mu + \lambda(v)) dv) \int_0^t \exp(-\int_0^u (\mu + \lambda(v)) dv) du > 0$  for all  $t \geq 0$ . Likewise, we can show in a similar manner that  $V_k(t) > 0, E_k > 0, A_k(t) > 0, I_k(t) > 0,$  and  $R_k(t) > 0$  for all  $k = 1, 2, \dots, n, t \geq 0$ . The result follows since  $dN/dt = \mu - \mu N, N(t_0) = 1$ .

The long term behavior of the solutions of model (1) depends on certain thresholds of the reproduction number. The reproduction number and the thresholds are calculated in sections to come.

### 2.3 Reproduction number

Define

$$\begin{aligned} c_k &= (1 - p)a_k\beta_k + pw_k\gamma_k, \\ \bar{c}_k &= (1 - p)a_k\theta_k + pw_kr_k, \quad \text{for } k = 1, 2, \dots, n. \end{aligned} \tag{5}$$

The disease-free equilibrium, denoted  $\mathcal{E}_0$ , of the system (1) is given by

$$\mathcal{E}_0 = \{S^0 = 1 - q, V_1^0 = q_k, \dots, V_n^0 = q_n, E_1^0 = 0, \dots, E_n^0 = 0, A_1^0 = 0, \dots, A_n^0 = 0, I_1^0 = 0, \dots, I_n^0 = 0, R_1^0 = 0, \dots, R_n^0 = 0\}. \tag{6}$$

The dynamic model (1) can be written in the form

$$d\mathbf{x} = (\mathcal{F}(\mathbf{x}) - \mathcal{V}(\mathbf{x})) dt, \tag{7}$$

where

$$\mathbf{x} = (E_1, \dots, E_n, A_1, \dots, A_n, I_1, \dots, I_n, R_1, \dots, R_n, V_1, \dots, V_n, S)^T, \tag{8}$$

$\mathcal{F}_j$  denotes the rate of appearance of new infections in compartment  $j$ ,  $\mathcal{V}_j = \mathcal{V}_j^- - \mathcal{V}_j^+$  with  $\mathcal{V}_j^+$  and  $\mathcal{V}_j^-$  denoting the rate of transfer of individuals in and out of compartment  $j$ , respectively [52]. For any vector  $\mathbf{u} = (u_1, \dots, u_n)^T$ , define the  $n \times n$  matrix  $\mathbf{M}(\mathbf{u})$  by the diagonal matrix

$$\mathbf{M}(\mathbf{u}) = \text{diag}(u_1, u_2, \dots, u_n). \tag{9}$$

Define

$$Q_k = \sum_{j=1}^k q_{j-1}, \tag{10}$$

where  $q_0 = 0$ . Let  $\tilde{\beta}$  and  $\tilde{\gamma}$  be two vectors with entries

$$\begin{aligned} \tilde{\beta}_k &= (1 - q + Q_k)\beta_k, \\ \tilde{\gamma}_k &= (1 - q + Q_k)\gamma_k, \end{aligned}$$

respectively, where  $\tilde{\beta}_1 = (1 - q)\beta_1$  and  $\tilde{\gamma}_1 = (1 - q)\gamma_1$ . Define  $\mathbf{F}$  and  $\mathbf{V}$  such that

$$\begin{aligned} (\mathbf{F}_{ij}) &= \left( \frac{\partial \mathcal{F}_i}{\partial \mathbf{x}_j}(\mathcal{E}_0) \right), \\ (\mathbf{V}_{ij}) &= \left( \frac{\partial \mathcal{V}_i}{\partial \mathbf{x}_j}(\mathcal{E}_0) \right), \end{aligned}$$

where  $(i, j)$  with  $1 \leq i, j \leq 3n$  corresponds to the index of the infected compartments. Based on our model (1), there are three infected compartments  $E, A$ , and  $I$ , each assumed to have  $n$ -different strains so that the infected compartments are the first  $3n$  entries of  $\mathbf{x}$  in (8). Let  $\mathbf{0}_{n \times n}$  represents the zero square matrix of order  $n$ . The corresponding matrices  $F$  and  $V$  are calculated as

$$\begin{aligned} \mathbf{F} &= \begin{pmatrix} \mathbf{0}_{n \times n} & \mathbf{M}(\tilde{\gamma}) & \mathbf{M}(\tilde{\beta}) \\ \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \end{pmatrix}, \\ \mathbf{V} &= \begin{pmatrix} \mathbf{M}(\mathbf{e}) & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ -p\mathbf{M}(\lambda) & \mathbf{M}(\mathbf{a}) & \mathbf{0}_{n \times n} \\ -(1-p)\mathbf{M}(\lambda) & \mathbf{0}_{n \times n} & \mathbf{M}(\mathbf{w}) \end{pmatrix}, \end{aligned}$$



where  $M(\mathbf{u})$  is as defined in (9) and  $\tilde{\beta}, \tilde{\gamma}, \mathbf{e}, \lambda, \mathbf{a}$ , and  $\mathbf{w}$  are vectors. From the matrix

$$V^{-1} = \begin{pmatrix} M(\mathbf{e})^{-1} & \mathbf{0}_{n \times n} & \mathbf{0}_{n \times n} \\ pM(\mathbf{e})^{-1}M(\mathbf{a})^{-1}M(\lambda) & M(\mathbf{a})^{-1} & \mathbf{0}_{n \times n} \\ (1-p)M(\mathbf{e})^{-1}M(\mathbf{w})^{-1}M(\lambda) & \mathbf{0}_{n \times n} & M(\mathbf{w})^{-1} \end{pmatrix},$$

where  $M(\mathbf{e})^{-1}$  is the diagonal matrix  $M(\mathbf{e})^{-1} = \text{diag}\left(\frac{1}{e_1}, \frac{1}{e_2}, \dots, \frac{1}{e_n}\right)$ , we see that the average length of time an individual spent being exposed, asymptotically infected, symptomatically infected with strain  $k$  is  $1/e_k, 1/a_k$ , and  $1/w_k$ , respectively. Also, the average length of time an individual exposed to strain  $k$  spent being asymptomatic with the same strain during its life time is expected to be  $\frac{pw_k\lambda_k}{a_k e_k w_k}$ . The average length of time an individual exposed to strain  $k$  spent being symptomatic with strain  $k$  during its life time is expected to be  $\frac{(1-p)a_k\lambda_k}{a_k e_k w_k}$ . Since the transmission rates of symptomatic and asymptomatic individuals with strain  $k$  in a susceptible population are  $\beta_k$  and  $\gamma_k$ , respectively, it follows that at the emergence of a new strain  $k$  ( $k \geq 2$ ), the expected number of new infections produced by individual with such strain in a completely susceptible non-vaccinated population, and a population completely vaccinated against strain  $j$ , are  $\frac{(1-q)((1-p)a_k\beta_k + pw_k\gamma_k)\lambda_k}{a_k e_k w_k} = \frac{(1-q)c_k\lambda_k}{a_k e_k w_k}$  and  $\frac{q_{j-1}((1-p)a_k\beta_k + pw_k\gamma_k)\lambda_k}{a_k e_k w_k} = \frac{q_{j-1}c_k\lambda_k}{a_k e_k w_k}$ , respectively,  $j = 1, 2, \dots, n$ , where we set  $q_0 = 0$ .

The expected number,  $\mathcal{R}_k$ , of new infections produced by individual with such strain  $k$  in a population containing susceptible and vaccinated individuals is obtained as

$$\mathcal{R}_k = \frac{(1-q)c_k\lambda_k + \sum_{j=1}^k q_{j-1}c_k\lambda_k}{a_k e_k w_k} = \frac{(1-q + Q_k)c_k\lambda_k}{a_k e_k w_k}, \quad k = 1, 2, \dots, n, \tag{11}$$

where  $1 - q$  is the proportion of those that are susceptible but not vaccinated. Using the next generation matrix [52], since  $\mathbf{F}$  is non-negative and  $\mathbf{V}$  is a non-singular  $M$ -matrix [53], the number  $\mathcal{R}_k$  obtained in (11) is the  $(k, k)$ -th entry of the next generation matrix  $\mathbf{FV}^{-1}$ , for  $k = 1, 2, \dots, n$ . It is the expected number of new infections in compartment  $k$  (compartment exposed to strain  $k$ ) produced by infected individual originally introduced into the same compartment.

**Remark 2. Effect of Vaccination**

We remark here that the expected number  $\mathcal{R}_k$  of new infections in the compartment exposed to strain  $k$  depends on the vaccination rates  $q_l, l = k, k + 1, \dots, n$ . If no one is vaccinated in the population (that is, if  $q_l = 0$  for all  $l = 1, 2, \dots, n$ ), then the expected number of infections caused by strain  $k$  is obtained to be

$$\mathcal{R}_{c,k} = \frac{c_k\lambda_k}{a_k e_k w_k}, \quad k = 1, 2, \dots, n. \tag{12}$$

It follows from Remark 1 that this number is clearly more than the reproduction number obtained in (11) where some individuals are receiving vaccination in the population. The reproduction number  $\mathcal{R}_k$  can be written in terms  $\mathcal{R}_{c,k}$  as

$$\mathcal{R}_k = (1 - q + Q_k)\mathcal{R}_{c,k}, \tag{13}$$

showing that the ratio of numbers of infection caused by strain  $k$  in a population with vaccination to a population without vaccination is  $1 - q + Q_k : 1$ . If individuals are not vaccinated against strain  $k$  (that is, if  $q_k = 0$  for fixed  $k$ ), then the expected number of infections caused by

strain  $k$  increases to

$$\frac{\left(1 - \sum_{\substack{j=1 \\ j \neq k}}^n q_j\right) c_k \lambda_k + \sum_{j=1}^k q_{j-1} c_k \lambda_k}{a_k e_k w_k} = \frac{\left(1 - \sum_{j=k+1}^n q_j\right) c_k \lambda_k}{a_k e_k w_k},$$

which is the same as  $\mathcal{R}_k + \frac{q_k c_k \lambda_k}{a_k e_k w_k}$ . This shows that a non vaccinated strain  $k$  infected individual produces  $\frac{q_k c_k}{(1-q)c_k + \sum_{j=1}^{k-1} q_j c_k} \times 100\%$  more infections than a vaccinated strain  $k$  infected individual. If

on the other hand, everyone in the population is vaccinated (that is,  $q = 1$ ), then the expected number of infections caused by strain  $k$  significantly reduced to

$$\frac{\sum_{j=1}^k q_{j-1} c_k \lambda_k}{a_k e_k w_k}, \quad k = 1, 2, \dots, n.$$

That is, a completely vaccinated population produces  $\frac{1-q}{1-q+Q_k} \times 100\%$  lesser infections than population not completely vaccinated. These analyses show the importance of being vaccinated in the population.

The reproduction number  $\mathcal{R}_0$  for the system, which is defined as the expected number of secondary infections produced by a typical infected individuals over the course of its infectious period, is calculated using the next generation matrix approach by van den Driessche et al. [52] as

$$\mathcal{R}_0 = \max_{1 \leq k \leq n} \left\{ (1 - q + Q_k) \frac{c_k \lambda_k}{a_k e_k w_k} \right\}. \tag{14}$$

We shall later show that if  $\mathcal{R}_0 < 1$ , then we expect a typical infected individual to produce less than one new infected individual, meaning all strains of the disease will eventually die out in the population (in the presence of vaccination). Although it is well known that a value of the reproduction number greater than one means that epidemic will persist in the population [8, 52], we shall later show that a strain with a reproduction number greater than 1 can still die out on the long run if a newer emerging strain has a greater reproduction number. The value  $\mathcal{R}_k$  on the other hand, can be interpreted as the reproduction number for typical individual infected with strain  $k$  of the disease, for  $k = 1, 2, \dots, n$ .

**Remark 3.** In the case where  $\beta_k \neq \bar{\beta}_{k,j}$  and  $\gamma_k \neq \bar{\gamma}_{k,j}$  such that  $\beta_k > \bar{\beta}_{k,j}$  and  $\gamma_k > \bar{\gamma}_{k,j}$  for all  $1 \leq k, j \leq n$  due to some form of partial immunity as a result of vaccines, then the strain  $k$  reproduction number  $\mathcal{R}_k$  and the reproduction number  $\mathcal{R}_0$  in (11) and (14), respectively, are obtained as

$$\mathcal{R}_k = \frac{(1 - q) c_k \lambda_k + \sum_{j=1}^k q_{j-1} \hat{c}_{k,j-1} \lambda_k}{a_k e_k w_k}, \quad k = 1, 2, \dots, n,$$

$$\mathcal{R}_0 = \max_{1 \leq k \leq n} \left\{ \frac{(1 - q) c_k \lambda_k + \sum_{j=1}^k q_{j-1} \hat{c}_{k,j-1} \lambda_k}{a_k e_k w_k} \right\}, \tag{15}$$

where

$$\hat{c}_{k,j} = (1 - p)a_k\bar{\beta}_{k,j} + pw_k\bar{\gamma}_{k,j}, \quad j = 1, \dots, k - 1, \tag{16}$$

and we set  $\hat{c}_{k,0} = 0 \forall k = 1, 2, \dots, n$ . It is easy to show that  $c_k > \hat{c}_{k,j}$  in this case, so that the estimates of the reproduction numbers given in (15) are greater than those given in (11) and (14), respectively. This shows that the reproduction number reduces as vaccines reduce the infectivity of the viruses (as expected).

**Remark 4.** The reproduction number  $\mathcal{R}_0$  is for the case where some members of the population are vaccinated against certain strain of the disease, that is, case with compartments  $V_1, V_2, \dots, V_n$  in the population. As explained in Remark 2, the reproduction number, denoted  $\mathcal{R}_c$ , for a completely susceptible non-vaccinated population is given by

$$\mathcal{R}_c = \max_{1 \leq k \leq n} \left\{ \frac{c_k \lambda_k}{a_k e_k w_k} \right\}. \tag{17}$$

### 2.4 Existence of equilibrium points

We discuss conditions under which equilibrium points of (1) exist. System (1) has many equilibrium points. Let  $P = \{1, 2, \dots, n\}$  denotes set of indices representing order of existence of new strain in the population, with  $2^P$  denoting the power set of  $P$ . Let  $\mathcal{S}_r \in 2^P$  denotes a subset of  $2^P$  with  $r$  number of strains,  $r = 0, 1, \dots, n$ , with  $r = 0$  representing disease-free case. We study the existence conditions for the equilibrium point corresponding to the scenario where only strains in  $\mathcal{S}_r$  survives,  $r = 0, 1, \dots, n$ . The case  $r = 0$  exists for the disease-free equilibrium  $\mathcal{E}_0$  case. The disease-free equilibrium is given in (6). For the case  $r = 1$  representing case where only one strain survives, we shall denote the strain by strain  $m, m = 1, 2, \dots, n$ , and the equilibrium referred to as the strain  $m$  equilibrium point and denoted  $\mathcal{E}_m$ . We study the conditions under which such equilibrium point exists, and in general, we also study conditions under which strains with indices in  $\mathcal{S}_r$  survives, for  $r = 1, 2, \dots, n$ .

**Theorem 2.** The strain  $m$  unique equilibrium point, denoted  $\mathcal{E}_m$ , for the epidemic model (1) exists in the feasible region  $\mathcal{T}$  provided  $\mathcal{R}_m > 1$ .

*Proof.* The strain  $m$  equilibrium point  $\mathcal{E}_m$  is obtained by solving the equation  $h(t) = 0$  and setting  $E_k = A_k = I_k = R_k = 0$  for  $1 \leq k \neq m \leq n$ , where  $h(t)$  is a vector in (4) containing the right hand side of (1). It follows from the equation governing  $A_m$  and  $I_m$  that  $A_m = \frac{p\lambda_m}{a_m} E_m$  and  $I_m = \frac{(1-p)\lambda_m}{w_m} E_m$ . Substituting these into the equations governing  $S, V_k, k = 1, 2, \dots, n, E_m$ , and  $R_m$ , we have

$$\begin{aligned} S &= \frac{(1 - q)\mu}{\mu + \frac{c_m \lambda_m}{a_m w_m} E_m}, \\ V_k &= \begin{cases} \frac{q_k \mu}{\mu + \frac{c_m \lambda_m}{a_m w_m} E_m}, & \text{if } k < m, \\ q_k, & \text{if } k \geq m, \end{cases} \\ E_m &= \frac{1}{c_m \lambda_m e_m} (\mu c_m \lambda_m (1 - q) + \mu c_m \lambda_m Q_m - \mu a_m w_m e_m), \\ R_m &= \frac{(1 - p)a_m \theta_m + p r_m w_m \lambda_m E_m}{\mu a_m w_m} \end{aligned}$$

where  $c_m$  and  $\bar{c}_m$  are defined in (5). If  $\mathcal{R}_m > 1$ , the strain  $m$  equilibrium point  $\mathcal{E}_m$  is obtained as

$$\mathcal{E}_m = \{S^*, V_1^*, \dots, V_n^*, E_1^*, \dots, E_n^*, A_1^*, \dots, A_n^*, I_1^*, \dots, I_n^*, R_1^*, \dots, R_n^*\}, \tag{18}$$

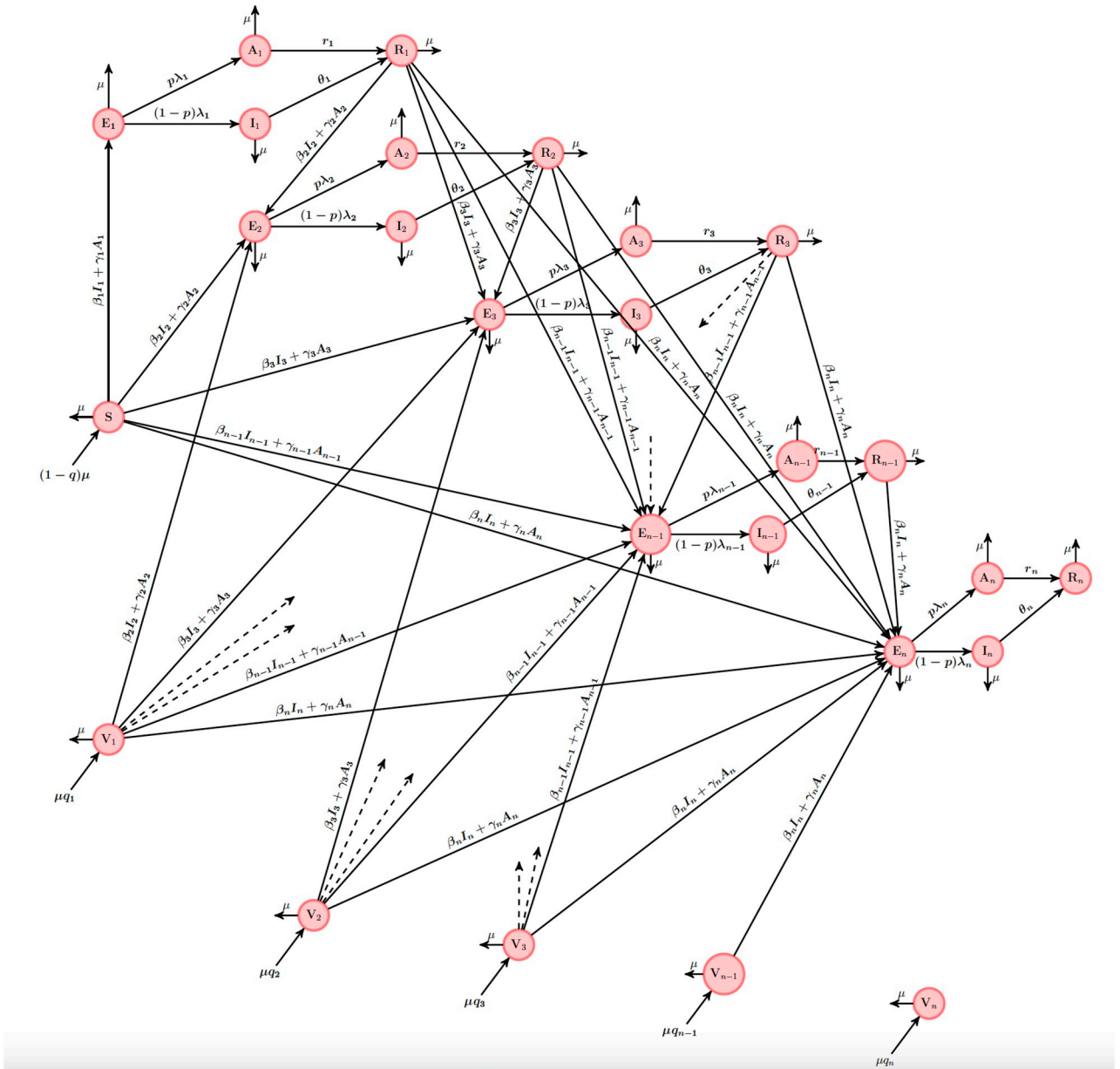
where

$$\begin{aligned} E_m^* &= \frac{\mu a_m w_m}{c_m \lambda_m} (\mathcal{R}_m - 1), \\ S^* &= \frac{(1 - q)}{\mathcal{R}_m} \\ V_k^* &= \begin{cases} \frac{q_k}{\mathcal{R}_m}, & \text{if } k < m, \\ q_k, & \text{if } k \geq m, \end{cases} \\ A_m^* &= \frac{p \lambda_m}{a_m} E_m^*, \\ I_m^* &= \frac{(1 - p) \lambda_m}{w_m} E_m^*, \\ R_m^* &= \frac{\bar{c}_m}{\mu a_m w_m} \lambda_m E_m^*, \end{aligned} \tag{19}$$

and  $E_k^* = A_k^* = I_k^* = R_k^* = 0$  for  $1 \leq k \neq m \leq n$ . We can show that  $S^*, \{V_k^*\}_{k=1}^n, E_m^*, A_m^*, I_m^*, R_m^* \in \mathcal{T}$  by using (5) to show that

$$\begin{aligned} S^* + \sum_{k=1}^n V_k^* + E_m^* + A_m^* + I_m^* + R_m^* &= \frac{(1 - q)\mu}{\mu + \frac{c_m \lambda_m}{a_m w_m} E_m^*} + \sum_{k=1}^{m-1} \frac{q_k \mu}{\mu + \frac{c_m \lambda_m}{a_m w_m} E_m^*} + \sum_{k=m}^n q_k + E_m^* \\ &\quad + \frac{p w_m + (1 - p) a_m \lambda_m E_m^*}{a_m w_m} + \frac{p w_m r_m + (1 - p) a_m \theta_m \lambda_m E_m^*}{\mu a_m w_m} \\ &= (1 - q) + \sum_{k=1}^n q_k + E_m^* + \frac{p w_m + (1 - p) a_m \lambda_m E_m^*}{a_m w_m} \\ &\quad + \frac{p w_m r_m + (1 - p) a_m \theta_m \lambda_m E_m^*}{\mu a_m w_m} - \frac{e_m}{\mu} E_m^* \\ &= 1 + \frac{\lambda_m}{\mu} E_m^* + E_m^* - \frac{e_m}{\mu} E_m^* \\ &= 1. \end{aligned}$$

**Remark 5.** Theorem 2 shows that strain  $m$  alone persists at equilibrium if the expected number of secondary infections produced by strain  $m$  infected individual is greater than one. We shall later show in Theorem 8 that this equilibrium point is stable globally if  $\mathcal{R}_k < 1$  for  $1 \leq k \neq m \leq n$  (guaranteeing  $E_k^* = A_k^* = I_k^* = R_k^* = 0$  for  $1 \leq k \neq m \leq n$ ) and  $\mathcal{R}_m > 1$ . Theorem 2 is also valid for the case where  $\beta_k > \bar{\beta}_{k,j}$  and  $\gamma_k > \bar{\gamma}_{k,j}$ ,  $j = 1, 2, \dots, k - 1$ . The proof of this is shown in Theorems 17 and 18.



**Fig 2. Schematic diagram for the epidemic model (35).** The circle compartments represent group of individuals.

<https://doi.org/10.1371/journal.pone.0271446.g002>

Suppose  $\mathcal{S}_2 = \{\tau_1, \tau_2\}$  with strain  $1 \leq \tau_1 < \tau_2 \leq n, \tau_j \in \mathbb{Z}_+ \cup \{0\}, j = 1, \dots, 2$ . We denote the equilibrium point corresponding to the case where strains  $\tau_1, \tau_2$  survive by  $\mathcal{E}_{\mathcal{S}_2}$ . We give theorem under which the equilibrium  $\mathcal{E}_{\mathcal{S}_2}$  exists.

**Theorem 3.** *The epidemic model (1) has an equilibrium point  $\mathcal{E}_{\mathcal{S}_2}$  in the feasible region  $\mathcal{I}$  provided*

$$\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + \left(Q_{\tau_2} - Q_{\tau_1}\right) \frac{\mathcal{R}_{\tau_1}}{1 - q + Q_{\tau_1}}} \tag{20}$$

*Proof.* Condition (20) implies that  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1$ . The equilibrium point  $\mathcal{E}_{S_2}$  is obtained as

$$\mathcal{E}_{S_2} = \{S^+, V_1^+, \dots, V_n^+, E_1^+, \dots, E_n^+, A_1^+, \dots, A_n^+, I_1^+, \dots, I_n^+, R_1^+, \dots, R_n^+\}, \tag{21}$$

where

$$\begin{aligned} E_{\tau_1}^+ &= \frac{\mu}{e_{\tau_1}} \frac{\mathcal{R}_{\tau_1} - \mathcal{R}_{\tau_2}}{\frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}}}, \\ E_{\tau_2}^+ &= \frac{\mu}{e_{\tau_2}} \left( \frac{1-q+Q_{\tau_2}}{\mathcal{R}_2} \right) \frac{\frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}} + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}} - \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}}{\frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}}}, \\ S^+ &= \frac{(1-q)\mu}{\mu + \sum_{j=1}^2 \left( \frac{c_{\tau_j} \lambda_{\tau_j}}{a_{\tau_j} w_{\tau_j}} E_{\tau_j}^+ \right)} = \frac{1-q}{\mathcal{R}_{\tau_1}}, \\ V_k^+ &= \begin{cases} \frac{q_k \mu}{\mu + \sum_{j=1}^2 \left( \frac{c_{\tau_j} \lambda_{\tau_j}}{a_{\tau_j} w_{\tau_j}} E_{\tau_j}^+ \right)}, & \text{if } k \leq \tau_1 - 1, \\ \frac{q_k \mu}{\mu + \frac{c_{\tau_2} \lambda_{\tau_2}}{a_{\tau_2} w_{\tau_2}} E_{\tau_2}^+}, & \text{if } \tau_1 \leq k < \tau_2, \\ q_k, & \text{if } k \geq \tau_2, \end{cases} \\ A_{\tau_k}^+ &= \frac{p \lambda_{\tau_k}}{a_{\tau_k}} E_{\tau_k}^+, \quad k = 1, 2, \\ I_{\tau_k}^+ &= \frac{(1-p) \lambda_{\tau_k}}{w_{\tau_k}} E_{\tau_k}^+, \quad k = 1, 2, \\ R_{\tau_k}^+ &= \frac{\bar{c}_{\tau_k} \lambda_{\tau_k}}{\mu a_{\tau_k} w_{\tau_k}} E_{\tau_k}^+, \quad k = 1, 2, \end{aligned} \tag{22}$$

and  $E_k^+ = A_k^+ = I_k^+ = R_k^+ = 0 \forall k \neq \tau_1, \tau_2$  if  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}}$ . If

$\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}}$ , then

$$\begin{aligned} E_{\tau_1}^+ &= \frac{\mu}{e_{\tau_1}} \frac{\mathcal{R}_{\tau_1} - \mathcal{R}_{\tau_2}}{\frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}}} > 0, \\ E_{\tau_2}^+ &= \frac{\mu a_{\tau_2} w_{\tau_2}}{c_{\tau_2} \lambda_{\tau_2}} \frac{\frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}} + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}} - \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}}{\frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}}} \\ &= \frac{\mu}{e_{\tau_2}} \frac{1}{\mathcal{R}_{\tau_2}} \frac{1}{\frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1-q+Q_{\tau_2}}} \left( \frac{1 + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}}{1-q+Q_{\tau_2}} \right) \left( \mathcal{R}_{\tau_2} - 1 - \frac{\mathcal{R}_{\tau_1} - 1}{1 + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1-q+Q_{\tau_1}}} \right) > 0. \end{aligned}$$

The result follows.

**Remark 6** Since  $\mathcal{R}_{\tau_1}$  is non-negative, condition (20) implies that  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1$ . This seems to suggest that the system is already in endemic state with strain  $\tau_1$  before the emergence of strain  $\tau_2$  caused an endemic. This is also confirmed in the work of Fudolig et al [10] where an SVIR model is used to analyze the transmission of the multistrains of the COVID-19 virus. Condition (20) is equivalent to

$$\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > \frac{b\mathcal{R}_{\tau_1}}{a + (b - a)\mathcal{R}_{\tau_1}}, \tag{23}$$

where  $a = 1 - q + Q_{\tau_1}$  and  $b = 1 - q + Q_{\tau_2}$ . We analyze the condition geometrically in Appendix B in S1 Appendix. It shows that for only strains  $\tau_1$  and  $\tau_2$  to remain in the system on the long run, the number of infection  $\mathcal{R}_{\tau_2}$  produced by strain  $\tau_2$ -infected individuals must be more than  $\frac{b\mathcal{R}_{\tau_1}}{a+(b-a)\mathcal{R}_{\tau_1}}$  but not up to the number  $\mathcal{R}_{\tau_1}$  produced by strain  $\tau_1$ -infected individuals. Once  $\mathcal{R}_{\tau_2}$  falls outside this region, then only strain  $\tau_1$  or  $\tau_2$  remains in the system on the long run. For instance, if  $\mathcal{R}_{\tau_2} < \frac{b\mathcal{R}_{\tau_1}}{a+(b-a)\mathcal{R}_{\tau_1}} < \mathcal{R}_{\tau_1}$ , it follows from (22) that the compartmental equilibrium value  $E_{\tau_1}^+ > 0$  and  $E_{\tau_2}^+ < 0$ , so that the values  $I_{\tau_1}^+, A_{\tau_1}^+$ , and  $R_{\tau_1}^+$  are positive but  $I_{\tau_2}^+, A_{\tau_2}^+$ , and  $R_{\tau_2}^+$  are negative, showing that only strain  $\tau_1$  remains on the long run. In the same sense, we can show that only one strain remains in the system on the long run if  $\mathcal{R}_{\tau_2} > \mathcal{R}_{\tau_1} > 1$ . This endemic region is shown graphically in Appendix B in S1 Appendix.

Also, it can be shown that

$$S^+ + \sum_{k=1}^n V_k^+ + \sum_{j=1}^2 (E_j^+ + A_j^+ + I_j^+ + R_j^+) = 1.$$

Suppose  $\mathcal{S}_r = \{\tau_1, \tau_2, \dots, \tau_r\}$  with strain  $1 \leq \tau_{j-1} < \tau_j \leq n, j = 2, \dots, r$ , and  $\tau_1, \tau_2, \dots, \tau_r \in \mathbb{Z}_+$ . We denote the equilibrium point corresponding to the case where strains  $\tau_1, \tau_2, \dots, \tau_r$  survive by  $\mathcal{E}_{\mathcal{S}_r}$ . We give theorem under which the equilibrium  $\mathcal{E}_{\mathcal{S}_r}, r = 2, \dots, n$ , exists.

**Theorem 4.** The epidemic model (1) has an equilibrium point  $\mathcal{E}_{\mathcal{S}_r}$  in the feasible region  $\mathcal{T}$  provided

$$\begin{aligned} \mathcal{R}_{\tau_{k-1}} > \mathcal{R}_{\tau_k} > \frac{(1 - q + Q_{\tau_k})(Q_{\tau_{k+1}} - Q_{\tau_{k-1}}) \frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}} \frac{\mathcal{R}_{\tau_{k+1}}}{1 - q + Q_{\tau_{k+1}}}}{(Q_{\tau_{k+1}} - Q_{\tau_k}) \frac{\mathcal{R}_{\tau_{k+1}}}{1 - q + Q_{\tau_{k+1}}} + (Q_{\tau_k} - Q_{\tau_{k-1}}) \frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}}}, \text{ for } k = 2, 3, \dots, r - 1, \\ \mathcal{R}_{\tau_r} > 1 + \frac{\mathcal{R}_{\tau_{r-1}} - 1}{1 + (Q_{\tau_r} - Q_{\tau_{r-1}}) \frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}}}. \end{aligned} \tag{24}$$

*Proof.* Condition (24) implies that  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} \dots > \mathcal{R}_{\tau_r} > 1$ . The equilibrium point  $\mathcal{E}_{\mathcal{S}_r}$  is obtained as

$$\mathcal{E}_{\mathcal{S}_r} = \{S^+, V_1^+, \dots, V_n^+, E_1^+, \dots, E_n^+, A_1^+, \dots, A_n^+, I_1^+, \dots, I_n^+, R_1^+, \dots, R_n^+\}, \tag{25}$$

where

$$\begin{aligned}
 E_{\tau_1}^+ &= \frac{\mu}{e_{\tau_1}} \frac{\mathcal{R}_{\tau_1} - \mathcal{R}_{\tau_2}}{\mathcal{R}_{\tau_1} - 1 - q + Q_{\tau_1} - \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}}}, \\
 E_k^+ &= \frac{\mu}{e_{\tau_k}} \frac{(Q_k - Q_{\tau_{k-1}}) \frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}} \frac{\mathcal{R}_{\tau_k}}{1 - q + Q_{\tau_k}} - (Q_{\tau_{k+1}} - Q_{\tau_{k-1}}) \frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}} \frac{\mathcal{R}_{\tau_{k+1}}}{1 - q + Q_{\tau_{k+1}}} + (Q_{\tau_{k+1}} - Q_{\tau_k}) \frac{\mathcal{R}_{\tau_k}}{1 - q + Q_{\tau_k}} \frac{\mathcal{R}_{\tau_{k+1}}}{1 - q + Q_{\tau_{k+1}}}}{\left( \frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}} - \frac{\mathcal{R}_{\tau_k}}{1 - q + Q_{\tau_k}} \right) \left( \frac{\mathcal{R}_{\tau_k}}{1 - q + Q_{\tau_k}} - \frac{\mathcal{R}_{\tau_{k+1}}}{1 - q + Q_{\tau_{k+1}}} \right)}, \quad k = 2, 3, \dots, r - 1, \\
 E_{\tau_r}^+ &= \frac{\mu a_{\tau_r} w_{\tau_r}}{c_{\tau_r} \lambda_{\tau_r}} \frac{\mathcal{R}_{\tau_r}}{1 - q + Q_{\tau_r}} + (Q_r - Q_{\tau_{r-1}}) \frac{\mathcal{R}_{\tau_r}}{1 - q + Q_{\tau_r}} \frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}} - \frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}}, \\
 S^+ &= \frac{(1 - q)\mu}{\mu + \sum_{j=1}^r \left( \frac{c_j \lambda_{\tau_j}}{a_j w_j} E_j^+ \right)} \tag{26} \\
 V_k^+ &= \begin{cases} \frac{q_k \mu}{\mu + \sum_{j=1}^r \left( \frac{c_j \lambda_{\tau_j}}{a_j w_j} E_j^+ \right)}, & \text{if } \tau_{l-1} \leq k \leq \tau_l - 1, \quad l = 1, 2, \dots, r, \quad \tau_0 = 1, \\ q_k, & \text{if } \tau_r \leq k \leq n, \end{cases} \\
 A_{\tau_l}^+ &= \frac{\rho \lambda_{\tau_l}}{a_{\tau_l}} E_{\tau_l}^+, \quad l = 1, 2, \dots, r, \\
 I_{\tau_l}^+ &= \frac{(1 - \rho) \lambda_{\tau_l}}{w_{\tau_l}} E_{\tau_l}^+, \quad l = 1, 2, \dots, r \\
 R_{\tau_l}^+ &= \frac{\bar{c}_{\tau_l} \lambda_{\tau_l}}{\mu a_{\tau_l} w_{\tau_l}} E_{\tau_l}^+, \quad l = 1, 2, \dots, r
 \end{aligned}$$

and  $E_k^+ = A_k^+ = I_k^+ = R_k^+ = 0$  for  $1 \leq k \neq \tau_1, \tau_2, \dots, \tau_r \leq n$  if condition (24) is satisfied. If condition (24) is satisfied, then

$$\begin{aligned}
 E_{\tau_1}^+ &> 0, \\
 E_{\tau_r}^+ &= \frac{\mu}{e_{\tau_r}} \frac{1}{\mathcal{R}_{\tau_r}} \frac{1}{\frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}} - \frac{\mathcal{R}_{\tau_r}}{1 - q + Q_{\tau_r}}} \left( \frac{1 + (Q_{\tau_r} - Q_{\tau_{r-1}}) \frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}}}{1 - q + Q_{\tau_r}} \right) \left( \mathcal{R}_{\tau_r} - 1 - \frac{\mathcal{R}_{\tau_{r-1}} - 1}{1 + (Q_{\tau_r} - Q_{\tau_{r-1}}) \frac{\mathcal{R}_{\tau_{r-1}}}{1 - q + Q_{\tau_{r-1}}}} \right) > 0,
 \end{aligned}$$

and for  $k = 2, 3, \dots, r - 1$ ,  $\frac{\mathcal{R}_{\tau_{k-1}}}{1 - q + Q_{\tau_{k-1}}} > \frac{\mathcal{R}_{\tau_k}}{1 - q + Q_{\tau_k}}$ , and  $E_{\tau_k}^+ > 0$ .

**Remark 7.** Condition (24) implies

$$\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > \mathcal{R}_{\tau_3} > \dots > \mathcal{R}_{\tau_r} > 1.$$

That is, the system is already in endemic state with strain  $\tau_i$  before the emergence of strain  $\tau_{i+1}$ ,  $i = 1, 2, \dots, r - 1$ , caused an endemic. Theorem 4 shows that if only strains  $\tau_1, \tau_2, \dots, \tau_r$  survive, then the system (without these strains) must converge to the disease-free equilibrium state.

**2.4.1 Endemic equilibrium.** The result for the endemic equilibrium can be calculated from Theorem 4 by extending the set  $\mathcal{S}_r$  to  $\mathcal{S}_n$ . We state the result without proof in the next theorem.



**Theorem 5.** *The epidemic model (1) has an endemic equilibrium point  $\mathcal{E}_{S_n}$  given by*

$$\begin{aligned}
 E_1^+ &= \frac{\mu}{e_1} \frac{\mathcal{R}_1 - \mathcal{R}_2}{\mathcal{R}_1 - \frac{\mathcal{R}_2}{1-q+Q_2}}, \\
 E_k^+ &= \frac{\mu}{e_k} \frac{(Q_k - Q_{k-1}) \frac{\mathcal{R}_{k-1}}{1-q+Q_{k-1}} \frac{\mathcal{R}_k}{1-q+Q_k} - (Q_{k+1} - Q_k) \frac{\mathcal{R}_{k-1}}{1-q+Q_{k-1}} \frac{\mathcal{R}_{k+1}}{1-q+Q_{k+1}} + (Q_{k+1} - Q_k) \frac{\mathcal{R}_k}{1-q+Q_k} \frac{\mathcal{R}_{k+1}}{1-q+Q_{k+1}}}{\left(\frac{\mathcal{R}_{k-1}}{1-q+Q_{k-1}} - \frac{\mathcal{R}_k}{1-q+Q_k}\right) \left(\frac{\mathcal{R}_k}{1-q+Q_k} - \frac{\mathcal{R}_{k+1}}{1-q+Q_{k+1}}\right)}, \quad k = 2, 3, \dots, n-1, \\
 E_n^+ &= \frac{\mu a_n w_n \frac{\mathcal{R}_n}{1-q+Q_n} + (Q_n - Q_{n-1}) \frac{\mathcal{R}_n}{1-q+Q_n} \frac{\mathcal{R}_{n-1}}{1-q+Q_{n-1}} - \frac{\mathcal{R}_{n-1}}{1-q+Q_{n-1}}}{\frac{\mathcal{R}_{n-1}}{1-q+Q_{n-1}} - \frac{\mathcal{R}_n}{1-q+Q_n}}, \\
 S^+ &= \frac{(1-q)\mu}{\mu + \sum_{j=1}^n \left(\frac{c_j \lambda_j}{a_j w_j} E_j^+\right)} \\
 V_k^+ &= \begin{cases} \frac{q_k \mu}{\mu + \sum_{j=k+1}^n \left(\frac{c_j \lambda_j}{a_j w_j} E_j^+\right)}, & \text{if } j = 1, 2, \dots, n-1, \\ q_k, & \text{if } j = n, \end{cases} \\
 A_k^+ &= \frac{p \lambda_k}{a_k} E_k^+, \quad k = 1, 2, \dots, n, \\
 I_k^+ &= \frac{(1-p)\lambda_k}{w_k} E_k^+, \quad k = 1, 2, \dots, n \\
 R_k^+ &= \frac{\bar{e}_k \lambda_k}{\mu a_k w_k} E_k^+, \quad k = 1, 2, \dots, n,
 \end{aligned} \tag{27}$$

in the feasible region  $\mathcal{T}$  provided

$$\begin{aligned}
 \mathcal{R}_{\tau_{k-1}} &> \mathcal{R}_{\tau_k} > \frac{(1-q+Q_{\tau_k})(Q_{\tau_{k+1}} - Q_{\tau_{k-1}}) \frac{\mathcal{R}_{\tau_{k-1}}}{1-q+Q_{\tau_{k-1}}} \frac{\mathcal{R}_{\tau_{k+1}}}{1-q+Q_{\tau_{k+1}}}}{\left(Q_{\tau_{k+1}} - Q_{\tau_k}\right) \frac{\mathcal{R}_{\tau_{k+1}}}{1-q+Q_{\tau_{k+1}}} + \left(Q_{\tau_k} - Q_{\tau_{k-1}}\right) \frac{\mathcal{R}_{\tau_{k-1}}}{1-q+Q_{\tau_{k-1}}}}, \quad \text{for } k = 2, 3, \dots, n-1, \\
 \mathcal{R}_{\tau_n} &> 1 + \frac{\mathcal{R}_{\tau_{n-1}} - 1}{1 + \left(Q_{\tau_n} - Q_{\tau_{n-1}}\right) \frac{\mathcal{R}_{\tau_{n-1}}}{1-q+Q_{\tau_{n-1}}}}.
 \end{aligned} \tag{28}$$

In the next section, we discuss the convergence of the system  $\tilde{y}$  in (4). We study conditions under which all strain infections are eradicated in the population on the long run. We also discuss the condition under which certain strain of the disease persists.

### 3 Stability analysis

In this section, we discuss the convergence of system (1) under certain conditions. That is, we study condition(s) under which the system converges to disease-free  $\mathcal{E}_0$  or equilibria  $\mathcal{E}_{S_r}$ ,  $r = 1, 2, \dots, n$ .

### 3.1 Stability analysis of the disease-free equilibrium point

**Theorem 6.** *The disease-free equilibrium  $\mathcal{E}_0$  is locally asymptotically stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_0 < 1$ .*

Given the initial condition  $\tilde{\mathbf{y}}_0 \in \mathcal{T}$ , Theorem 6 shows that if the system  $\tilde{\mathbf{y}}$  satisfying (1) starts near the initial point  $\tilde{\mathbf{y}}_0$ , then the system converges to the equilibrium point  $\mathcal{E}_0$  (that is, all the strain of infections are eradicated) provided the threshold  $\mathcal{R}_0 < 1$ . We use the idea presented in [27] to prove Theorem 6.

*Proof.* Let  $\mathcal{I}_{n \times n}$  be the  $n \times n$  identity matrix,  $\mathbf{y} = \tilde{\mathbf{y}} - \mathcal{E}_0$ ,  $\boldsymbol{\beta} = \{\beta_1, \dots, \beta_n\}$ ,  $\boldsymbol{\gamma} = \{\gamma_1, \dots, \gamma_n\}$ , the matrices  $\mathcal{U}_q(\boldsymbol{\beta})$  and  $\mathcal{U}_q(\boldsymbol{\gamma})$  defined by

$$\mathcal{U}_q(\boldsymbol{\beta}) = \begin{pmatrix} 0 & q_1\bar{\beta}_{2,1} & q_1\bar{\beta}_{3,1} & q_1\bar{\beta}_{4,1} & \cdots & q_1\bar{\beta}_{n,1} \\ 0 & 0 & q_2\bar{\beta}_{3,2} & q_2\bar{\beta}_{4,2} & \cdots & q_2\bar{\beta}_{n,2} \\ 0 & 0 & 0 & q_3\bar{\beta}_{4,3} & \cdots & q_3\bar{\beta}_{n,3} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & q_{n-1}\bar{\beta}_{n,n-1} \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{U}_q(\boldsymbol{\gamma}) = \begin{pmatrix} 0 & q_1\bar{\gamma}_{2,1} & q_1\bar{\gamma}_{3,1} & q_1\bar{\gamma}_{4,1} & \cdots & q_1\bar{\gamma}_{n,1} \\ 0 & 0 & q_2\bar{\gamma}_{3,2} & q_2\bar{\gamma}_{4,2} & \cdots & q_2\bar{\gamma}_{n,2} \\ 0 & 0 & 0 & q_3\bar{\gamma}_{4,3} & \cdots & q_3\bar{\gamma}_{n,3} \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & q_{n-1}\bar{\gamma}_{n,n-1} \\ 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix}$$

so that

$$\begin{aligned} G &= (0_{1 \times n} \quad -(1-q)\mathcal{U}_q(\boldsymbol{\gamma}) \quad -(1-q)\mathcal{U}_q(\boldsymbol{\beta})), \\ H &= (\mathbf{0}_{n \times n} \quad r\mathcal{I}_{n \times n} \quad \theta\mathcal{I}_{n \times n}), \\ \mathcal{A} &= \begin{pmatrix} \mathbf{M}(\mathbf{e}) & -\mathbf{M}(\tilde{\boldsymbol{\gamma}}) & -\mathbf{M}(\tilde{\boldsymbol{\beta}}) \\ -p\mathbf{M}(\lambda) & \mathbf{M}(\mathbf{a}) & \mathbf{0}_{n \times n} \\ -(1-p)\mathbf{M}(\lambda) & \mathbf{0}_{n \times n} & \mathbf{M}(\mathbf{w}) \end{pmatrix}, \end{aligned}$$

where  $\mathbf{M}$  is defined in (9). The linearization of the model (1) at the equilibrium point  $\mathcal{E}_0$  is derived as

$$\mathbf{y}' = \mathbf{A}\mathbf{y}, \tag{29}$$

where

$$\mathbf{A} = \begin{pmatrix} -\mu & 0_{1 \times n} & H & 0_{1 \times n} \\ 0_{n \times 1} & -\mu\mathcal{I}_{n \times n} & G & \mathbf{0}_{n \times n} \\ 0_{3n \times 1} & \mathbf{0}_{3n \times n} & -\mathcal{A} & \mathbf{0}_{3n \times n} \\ 0_{n \times 1} & \mathbf{0}_{n \times n} & H & -\mu\mathcal{I}_{n \times n} \end{pmatrix}.$$

In order to show that the equilibrium point  $\mathcal{E}_0$  is locally stable, we need to show that the maximum real part,  $s(\mathbf{A})$ , of  $\mathbf{A}$  is negative. From the structure of the matrix  $\mathbf{A}$ , this reduces to showing that the maximum real part of the eigenvalues of matrix  $-\mathcal{A}$  is negative (or equivalently,

$s(\mathcal{A}) > 0$ ). The determinant  $\det(\mathcal{A})$  of the matrix  $\mathcal{A}$  is

$$\det(\mathcal{A}) = \prod_{k=1}^n (1 - \mathcal{R}_k).$$

If  $\mathcal{R}_0 < 1$ , then  $\det(\mathcal{A}) > 0$ . Matrix  $\mathcal{A}$  is a non-singular  $Z$  matrix that can be written in the form

$$\mathcal{A} = \mathcal{L}\mathcal{U}, \tag{30}$$

where  $\mathcal{L}$  and  $\mathcal{U}$  are lower and upper diagonal matrices, respectively, with positive diagonals obtained as

$$\mathcal{L}_{ij} = \frac{1}{\mathcal{D}_j} \begin{vmatrix} \mathcal{A}_{1,1} & \mathcal{A}_{1,2} & \cdots & \mathcal{A}_{1,j} \\ \mathcal{A}_{2,1} & \mathcal{A}_{2,2} & \cdots & \mathcal{A}_{2,j} \\ \vdots & \vdots & \dots & \vdots \\ \mathcal{A}_{j-1,1} & \mathcal{A}_{j-1,2} & \cdots & \mathcal{A}_{j-1,j} \\ \mathcal{A}_{i,1} & \mathcal{A}_{i,2} & \cdots & \mathcal{A}_{i,j} \end{vmatrix}, \text{ for } i \geq j \neq 1, \quad \mathcal{L}_{i,1} = \frac{|\mathcal{A}_{i,1}|}{\mathcal{D}_1} \text{ for } i = 1, 2, \dots, 3n, \text{ and } 0 \text{ elsewhere,}$$

$$\mathcal{U}_{ij} = \frac{1}{\mathcal{D}_{i-1}} \begin{vmatrix} \mathcal{A}_{1,1} & \cdots & \mathcal{A}_{1,i-1} & \mathcal{A}_{1,j} \\ \mathcal{A}_{2,1} & \cdots & \mathcal{A}_{2,i-1} & \mathcal{A}_{2,j} \\ \vdots & \vdots & \vdots & \vdots \\ \mathcal{A}_{i,1} & \cdots & \mathcal{A}_{i,i-1} & \mathcal{A}_{i,j} \end{vmatrix}, \text{ for } 1 \neq i \leq j, \quad \mathcal{U}_{1,j} = \mathcal{A}_{1,j}, \text{ for } j = 1, 2, \dots, 3n, \text{ and } 0 \text{ elsewhere,}$$

where  $\mathcal{D}_0 := 1$ , and  $\mathcal{D}_j = \begin{vmatrix} \mathcal{A}_{1,1} & \mathcal{A}_{1,2} & \cdots & \mathcal{A}_{1,j} \\ \mathcal{A}_{2,1} & \mathcal{A}_{2,2} & \cdots & \mathcal{A}_{2,j} \\ \vdots & \vdots & \dots & \vdots \\ \mathcal{A}_{j,1} & \mathcal{A}_{j,2} & \cdots & \mathcal{A}_{j,j} \end{vmatrix}$  for  $j = 1, 2, \dots, 3n$ , with

$$\mathcal{D}_j = \prod_{k=1}^j e_k, \text{ if } 1 \leq j \leq n,$$

$$\mathcal{D}_{n+j} = \left( \prod_{l=1}^n e_l \right) \prod_{k=1}^j \left[ a_k \left( 1 - \mathcal{R}_k + (1-p) \left( (1-q)\beta_k + \sum_{l=1}^{k-1} \bar{\beta}_{k,l} q_l \right) a_k \lambda_k \right) \right], \text{ if } 1 \leq j \leq n,$$

$$\mathcal{D}_{2n+j} = \prod_{l=1}^j (a_l e_l w_l (1 - \mathcal{R}_l)) \prod_{k=j+1}^n \left[ a_k e_k \left( 1 - \mathcal{R}_k + (1-p) \left( (1-q)\beta_k + \sum_{l=1}^{k-1} \bar{\beta}_{k,l} q_l \right) a_k \lambda_k \right) \right], \text{ if } 1 \leq j \leq n-1,$$

$$\mathcal{D}_{3n} = \prod_{l=1}^n a_l e_l w_l (1 - \mathcal{R}_l).$$

The diagonals  $\mathcal{L}_{jj} = 1$  and  $\mathcal{U}_{ij} = \mathcal{D}_j / \mathcal{D}_{j-1} > 0$  if  $\mathcal{R}_0 < 1$ . Therefore, the matrix  $\mathcal{A}$  is a non-singular  $M$  matrix, and hence, a  $P$ -matrix. It follows from Berman [54] and Plemmons [53] that the maximum real part of the eigenvalues of  $-\mathcal{A}$  is negative.

**Remark 8.** Theorem 6 shows that if the population starts sufficiently close to  $\bar{y}_0$ , then the system  $\bar{y}$  converges to  $\mathcal{E}_0$  if  $\mathcal{R}_0 < 1$ . That is, the proportion of the susceptible size converges to  $1 - q$ , the proportion of the size of the vaccinated class immune to strain  $k$  converges to the vaccination rate  $q_k$  of individuals with strain  $k$  of infection, and no infected class, hence no need of recovery on the long run if  $\mathcal{R}_0 < 1$ .

We prove in the next theorem that the population, irrespective of where it starts from, converges to the point  $\mathcal{E}_0$  if  $\mathcal{R}_0 \leq 1$ .

**Theorem 7.** The disease-free equilibrium  $\mathcal{E}_0$  is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_0 \leq 1$ .

*Proof.* Define the Lyapunov function  $\mathcal{V}$  by

$$\bar{\mathcal{V}} = \left( S - S^0 - S^0 \ln \frac{S}{S^0} \right) + \sum_{k=1}^n \bar{\phi}_k \left( V_k - V_k^0 - V_k^0 \ln \frac{V_k}{V_k^0} \right) + \sum_{k=1}^n (\bar{z}_k E_k + \bar{f}_k I_k + \bar{g}_k A_k)$$

where

$$\begin{aligned} \bar{z}_k &= \mathcal{R}_k, \\ \bar{\phi}_k &= 1, \\ \bar{f}_k &= \frac{1}{w_k} \left( S^0 \beta_k + \sum_{j=1}^k \bar{\phi}_{j-1} V_{j-1}^0 \beta_k \right) \\ \bar{g}_k &= \frac{1}{a_k} \left( S^0 \gamma_k + \sum_{j=1}^k \bar{\phi}_{j-1} V_{j-1}^0 \gamma_k \right), \quad \text{for all } k = 1, 2, \dots, n, \end{aligned}$$

where  $\bar{\phi}_0 = 0$ . Let  $s = S/S^0$ ,  $v_k = V_k/V_k^0$ . If  $\mathcal{R}_k \leq 1$ , it follows that

$$\begin{aligned} d\bar{\mathcal{V}}/dt &= (1 - q)\mu - \mu(S - S^0) - \mu(1 - q)S^0/S + \sum_{k=1}^n \bar{\phi}_k \mu(q_k + V_k^0) \\ &\quad + \sum_{j=1}^n (\bar{z}_j - 1)(\beta_j I_j + \gamma_j A_j) S - \sum_{k=1}^{n-1} \sum_{j=k+1}^n (\bar{\phi}_k - \bar{z}_k)(\beta_j I_j + \gamma_j A_j) V_k - \mu \sum_{k=1}^n \bar{\phi}_k (V_k + q_k V_k^0/V_k) \\ &\quad + \sum_{k=1}^n \left( S^0 \beta_k + \sum_{j=1}^k \bar{\phi}_{j-1} V_{j-1}^0 \beta_k - \bar{f}_k w_k \right) I_k + \sum_{k=1}^n \left( S^0 \gamma_k + \sum_{j=1}^k \bar{\phi}_{j-1} V_{j-1}^0 \gamma_k - \bar{g}_k a_k \right) A_k \\ &\quad + \sum_{k=1}^n ((1 - p)\bar{f}_k \lambda_k + p\bar{g}_k \lambda_k - \bar{z}_k e_k) E_k \\ &\leq -\mu S^0 \left( s + \frac{1}{s} - 2 \right) - \sum_{k=1}^n \mu \bar{\phi}_k q_k \left( v_k + \frac{1}{v_k} - 2 \right) \end{aligned}$$

where

$$\begin{aligned}
 S^0 \beta_k + \sum_{j=1}^{k-1} \bar{\phi}_j V_j^0 \bar{\beta}_{k,j} - \bar{f}_k w_k &= 0, \\
 S^0 \gamma_k + \sum_{j=1}^k \bar{\phi}_{j-1} V_{j-1}^0 \gamma_k - \bar{g}_k a_k &= 0, \\
 (1-p)\bar{f}_k \lambda_k + p\bar{g}_k \lambda_k - z_k e_k &= \frac{1}{a_k w_k} \left( (1-q)c_k \lambda_k + \sum_{j=1}^k \bar{\phi}_{j-1} (q_{j-1} c_k \lambda_k) - z_k a_k e_k w_k \right) \\
 &= e_k (\mathcal{R}_k - z_k) = 0,
 \end{aligned}$$

and

$$2(\mu S^0 + \sum_{k=1}^n \mu \bar{\phi}_k q_k) = (1-q)\mu + \mu S^0 + \sum_{k=1}^n \bar{\phi}_k \mu (q_k + V_k^0) = 2\mu$$

using (6). Using the fact that the arithmetic mean of a list of non-negative real numbers is greater than or equal to the geometric mean of the same list [55], it follows that if  $\mathcal{R}_0 \leq 1$ , then  $d\bar{V}/dt \leq 0$ . Also,  $d\bar{V}/dt = 0$  if  $S = S^0 = 1 - q$ ,  $I_k = E_k = R_k = 0$ ,  $V_k = V_k^0 = q_k$  (or equivalently, if  $s = 1$ ,  $I_k = E_k = R_k = 0$ ,  $v_k = 1$ ) for all  $k = 1, 2, \dots, n$ . Since  $\mathcal{E}_0$  is the largest invariant set in the subset of  $\mathcal{T}$  where  $d\bar{V}/dt = 0$ , its global stability follows by the LaSalle’s Invariance Principle [56]

**Remark 9.** *Theorem 7 shows that the susceptible class converges to a fraction  $1 - q$  of the entire population size (this is simply the population size without the vaccinated group), the vaccinated class immune to strain  $k$  converges to a fraction  $q_k$  of population size, and no infected class, hence no need of recovery on the long run if  $\mathcal{R}_0 \leq 1$ . This suggests the threshold for disease eradication is the number  $\mathcal{R}_0$ .*

### 3.2 Bound for the critical vaccination threshold

Herd immunity is a state where significant proportion of the population is immune to an infection so that only few susceptible individuals can be infected and transmit the infection [57]. Classical vaccine-induced herd-immunity threshold suggests that the spread of a disease can be stopped by vaccinating certain fraction of the population. This might be invalid and biased due to many factors such as emergence of multi variant strains of the virus/disease, the dynamic nature of virus transmission, presence of immunity due to infection, changes in implementation and adherence to public health measures, and uncertainties in vaccine effectiveness and duration of immunity [57, 41]. These can cause the estimation of the Herd-immunity threshold to be imprecise. In this section, we aim to estimate a bound for the minimum proportion of the population that must be vaccinated in order for certain infection to die out in the population.

Let  $E_j \in (0, 1]$  denotes the proportion of vaccinated individuals against strain  $j$  who are protected by vaccines, for  $j = 1, 2, \dots, n$ . It follows from Remark 4 that we can write  $\mathcal{R}_k$  in (14) in terms of  $\mathcal{R}_{c,k}$  in (17) as follows:

$$\mathcal{R}_j = (1 - c_j E_j) \mathcal{R}_{c,j}, \tag{31}$$

for some constant  $c_j \in (0, 1)$ . If the vaccine is perfect and provides 100% immunity, then the vaccine effectiveness  $E_j = 1$ , otherwise,  $E_j < 1$  if it provides only partial immunity. According to Theorem 7, the population reaches herd immunity with respect to strain  $j$ , with incidence of

the infection declining, if  $\mathcal{R}_j \leq 1$ . This yield

$$c_j \geq \left(1 - \frac{1}{\mathcal{R}_{c,j}}\right) \frac{1}{E_j}.$$

The number

$$H_j = \left(1 - \frac{1}{\mathcal{R}_{c,j}}\right) \frac{1}{E_j} \tag{32}$$

is the minimum proportion of a population infected that must be vaccinated to stop strain  $j$  from spreading. This number is referred to as the herd immunity threshold [41]. In the presence of multiple vaccines for strains  $j = 1, 2, \dots, n$  with effectiveness  $E_1, E_2, \dots, E_n$ , respectively, let  $H_m$  and  $H_M$  denote the minimum  $\min_{1 \leq j \leq n} H_j$  and maximum  $\max_{1 \leq j \leq n} H_j$  herd immunity thresholds for the strains  $j = 1, 2, \dots, n$ , respectively. The interval

$$[H_m, H_M] \tag{33}$$

contains the maximum proportion of the population expected to be vaccinated to stop the spread of the disease.

**Remark 10.** *Caveats to the estimate*

We note here that the bound in (33) is calculated for a population satisfying the dynamics given in (1). The bound is expected to change in the emergence of a new strain of the virus. Also, the bound is calculated without taking into consideration the proportion of those who have immunity from the virus. These and more are some of the caveats to this estimates.

Sometimes new variants emerge and disappear. Other times, new variants persist. We study conditions under which new variant, say strain  $m$ , persists in the population. In the next section, we study the behavior of the system in the case where disease is not completely eradicated in the system.

### 3.3 Stability analysis of strain $m$ equilibrium

In the next theorem, we study how the population behaves on the long run if the number,  $\mathcal{R}_m$ , of new infections produced by infected individual with strain  $m$  is greater than one, while  $\mathcal{R}_k \leq 1$  for  $1 \leq k \neq m \leq n$ .

**Theorem 8.** *The strain  $m$  equilibrium  $\mathcal{E}_m$  for the epidemic model (1) is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_k \leq 1$  for all  $1 \leq k \neq m \leq n$  and  $\mathcal{R}_m > 1$ .*

*Proof.* According to Theorem 2, the strain  $m$  equilibrium  $\mathcal{E}_m$  exists and non-negative if  $\mathcal{R}_m > 1$ . Define the Lyapunov function  $\mathcal{V}$  by

$$\begin{aligned} \mathcal{V} = & \left( S - S^* - S^* \ln \frac{S}{S^*} \right) + z_m \left( E_m - E_m^* - E_m^* \ln \frac{E_m}{E_m^*} \right) + f_m \left( I_m - I_m^* - I_m^* \ln \frac{I_m}{I_m^*} \right) + g_m \left( A_m - A_m^* - A_m^* \ln \frac{A_m}{A_m^*} \right) \\ & + d_m \left( R_m - R_m^* - R_m^* \ln \frac{R_m}{R_m^*} \right) + \sum_{k=1}^n \varphi_k \left( V_k - V_k^* - V_k^* \ln \frac{V_k}{V_k^*} \right) + \sum_{\substack{k=1 \\ k \neq m}}^n (z_k E_k + f_k I_k + g_k A_k + d_k R_k) \end{aligned}$$

where

$$\begin{aligned} z_k &= \varphi_k = d_k = 1, \\ f_k &= \frac{1}{w_k} \left( S^* \beta_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^* \beta_k \right), \\ g_k &= \frac{1}{a_k} \left( S^* \gamma_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^* \gamma_k \right), \text{ for all } k = 1, 2, \dots, n, \end{aligned}$$

with  $\varphi_0 = 0$ . Let  $s = S/S^*$ ,  $v_k = V_k/V_k^*$ ,  $\Sigma_m = E_m/E_m^*$ ,  $\tau_m = I_m/I_m^*$ ,  $\Delta_m = A_m/A_m^*$ . It follows that

$$\begin{aligned} d\mathcal{V}/dt &= (1 - q)\mu - \mu(S - S^*) - \mu(1 - q)S^*/S + z_m e_m E_m^* + \sum_{k=1}^n \varphi_k \mu (q_k + V_k^*) + f_m w_m I_m^* + g_m a_m A_m^* \\ &+ \sum_{j=1}^n (z_j - 1)(\beta_j I_j + \gamma_j A_j) S - \sum_{k=1}^{n-1} \sum_{j=k+1}^n (\varphi_k - z_k)(\beta_j I_j + \gamma_j A_j) V_k - (1 - p) f_m \lambda_m I_m^* E_m / I_m \\ &- p g_m \lambda_m A_m^* E_m / A_m - z_m E_m^* ((\beta_m I_m / E_m + \gamma_m A_m / E_m) S + \sum_{k=1}^{m-1} V_k (\beta_m I_m / E_m + \gamma_m A_m / E_m)) \\ &- \mu \sum_{k=1}^n \varphi_k (V_k + q_k V_k^* / V_k) + \sum_{k=1}^n (S^* \beta_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^* \beta_k - f_k w_k) I_k + \sum_{k=1}^n (S^* \gamma_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^* \gamma_k - g_k a_k) A_k \\ &+ \sum_{k=1}^n ((1 - p) f_k \lambda_k + p g_k \lambda_k - z_k e_k) E_k \\ &\leq -\eta \left( s + \frac{1}{s} - 2 \right) - u \left( \frac{s \tau_m}{\Sigma_m} + \frac{\Sigma_m}{\tau_m} + \frac{1}{s} - 3 \right) - b \left( \frac{s \Delta_m}{\Sigma_m} + \frac{\Sigma_m}{\Delta_m} + \frac{1}{s} - 3 \right) - \sum_{k=1}^{m-1} \sigma_k \left( \frac{v_k \tau_m}{\Sigma_m} + \frac{\Sigma_m}{\tau_m} + \frac{1}{v_k} - 3 \right) \\ &- \sum_{k=1}^{m-1} \varphi_k \left( \frac{v_k \Delta_m}{\Sigma_m} + \frac{\Sigma_m}{\Delta_m} + \frac{1}{v_k} - 3 \right) - \sum_{k=1}^n \Upsilon_k \left( v_k + \frac{1}{v_k} - 2 \right) \end{aligned}$$

if  $\mathcal{R}_m > 1$  and  $\mathcal{R}_k \leq 1$  for  $1 \leq k \neq m \leq n$ , where

$$\begin{aligned} (1 - p) f_k \lambda_k + p g_k \lambda_k - z_k e_k &= \frac{1}{a_k w_k} \left( S^* c_k \lambda_k + \sum_{j=1}^{k-1} (\varphi_j V_j^* \bar{c}_{kj} \lambda_k) - a_k e_k w_k z_k \right) \\ &\begin{cases} = 0, & \text{if } k = m, \text{ using (19),} \\ \leq e_k (\mathcal{R}_k - 1) \leq 0, & \text{if } \mathcal{R}_k \leq 1, \text{ } k \neq m, \end{cases} \end{aligned}$$

$$\begin{aligned} \eta &= \mu S^*, \\ u &= z_m \beta_m I_m^* S^*, \\ b &= z_m \gamma_m A_m^* S^*, \\ \sigma_k &= z_m \beta_m I_m^* V_k^*, \text{ for } 1 \leq k \leq m - 1 \\ \varphi_k &= z_m \gamma_m A_m^* V_k^*, \text{ for } 1 \leq k \leq m - 1 \\ \Upsilon_k &= \mu \varphi_k V_k^*, \text{ for } 1 \leq k \leq n \end{aligned}$$

and

$$2(\eta + \sum_{k=1}^n \Upsilon_k) + 3(u + b + \sum_{k=1}^{m-1} (\sigma_k + \varphi_k)) = (1 - q)\mu + \mu S^* + f_m w_m I_m^* + g_m a_m A_m^* + z_m e_m E_m^* + \mu \sum_{j=1}^n \varphi_k (q_k + V_k^*)$$

using (19). Using the fact that the arithmetic mean of a list of non-negative real numbers is greater than or equal to the geometric mean of the same list [56], it follows that  $dV/dt \leq 0$ . Equality holds if  $S = S^*$ ,  $I_k = E_k = R_k = 0$  for all  $k \neq m$ ,  $V_k = V_k^*$  for all  $k = 1, 2, \dots, n$ ,  $I_m/I_m^* = E_m/E_m^* = R_m/R_m^* = 1$ . Since  $\mathcal{E}_m$  is the largest invariant set in the subset of  $\mathcal{T}$  where  $dV/dt = 0$ , its global stability follows by the LaSalle’s Invariance Principle [56].

**Remark 11.** Theorem 8 shows that the system  $\bar{y}$  converges to  $\mathcal{E}_m$  on the long run if  $\mathcal{R}_k \leq 1$  for  $1 \leq k \neq m \leq n$  and  $\mathcal{R}_m > 1$  irrespective of the starting point. That is, if the number,  $\mathcal{R}_m$ , of new infections is greater than one while  $\mathcal{R}_k \leq 1$  for  $1 \leq k \neq m \leq n$ , then on the long run, the susceptible class converges to a fraction  $S^* = \frac{(1-q)}{\mathcal{R}_m} < 1 - q$  of the entire population size, no exposure to any strain other than strain  $m$  in the population (hence no infection other than those caused by strain  $m$ , and no need for recovery), and before the existence of strain  $m$  (that is,  $k < m$ ), the vaccinated class immune to strain  $k < m$  converges to a fraction  $V_k^* = \frac{q_k}{\mathcal{R}_m} < q_k$  of the population size. We see from Theorem 7 that if there is no endemic in the population, the susceptible population  $S$  converges on the long run to  $(1 - q) \times 100\%$  of the population. This number reduces by  $(1 - \frac{1}{\mathcal{R}_m}) \times 100\%$  in the presence of an emerging strain  $m$ . Also, from Theorem 7, the vaccinated class  $V_k$  converges to  $q_k \times 100\%$  of the population size if there is no endemic in the population. This number also reduces by  $(1 - \frac{1}{\mathcal{R}_m}) \times 100\%$  in the presence of an emerging strain  $m$ . If  $k > m$  on the other hand, since  $\mathcal{R}_k \leq 1$  for  $1 \leq k \neq m \leq n$ , we have the vaccinated class immune to strain  $k$  ( $k > m$ ) converging to a fraction  $V_k^* = q_k$  of the population, no exposure to any strain other than strain  $m$  in the population (hence no infection and no need for recovery).

### 3.4 Stability analysis of the equilibrium point $\mathcal{E}_{S_2}$

We give the proof of the stability of the equilibrium point  $\mathcal{E}_{S_2}$  in the next theorem.

**Theorem 9.** The equilibrium point  $\mathcal{E}_{S_2}$  is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_k < 1$  for  $1 \leq k \neq \tau_1, \tau_2 \leq n$  and condition (20) is satisfied.

*Proof.* Assume  $\mathcal{R}_k \leq 1$  for all  $k \notin \tau_1, \tau_2$  and condition (20) is satisfied. The existence of the equilibrium point  $\mathcal{E}_{S_2}$  follows from Theorem 3. Define the Lyapunov function  $\mathcal{V}_2$  by

$$\begin{aligned} \mathcal{V}_2 = & \left( S - S^+ - S^+ \ln \frac{S}{S^+} \right) + \sum_{l=1}^2 z_{\tau_l} \left( E_{\tau_l} - E_{\tau_l}^+ - E_{\tau_l}^+ \ln \frac{E_{\tau_l}}{E_{\tau_l}^+} \right) + \sum_{l=1}^2 f_{\tau_l} \left( I_{\tau_l} - I_{\tau_l}^+ - I_{\tau_l}^+ \ln \frac{I_{\tau_l}}{I_{\tau_l}^+} \right) \\ & + \sum_{l=1}^2 g_{\tau_l} \left( A_{\tau_l} - A_{\tau_l}^+ - A_{\tau_l}^+ \ln \frac{A_{\tau_l}}{A_{\tau_l}^+} \right) + \sum_{k=1}^n \varphi_k \left( V_k - V_k^+ - V_k^+ \ln \frac{V_k}{V_k^+} \right) + \sum_{\substack{k=1 \\ k \neq \tau_1, \tau_2}}^n (z_k E_k + f_k I_k + g_k A_k + d_k R_k) \end{aligned}$$



where

$$\begin{aligned}
 z_k &= \varphi_k = d_k = 1, \\
 f_k &= \frac{1}{w_k} \left( S^+ \beta_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^+ \beta_k \right), \\
 g_k &= \frac{1}{a_k} \left( S^+ \gamma_k + \sum_{j=1}^k \varphi_{j-1} V_{j-1}^+ \gamma_k \right), \quad \text{for all } k = 1, 2, \dots, n,
 \end{aligned}
 \tag{34}$$

with  $\varphi_0 = 0$ . The derivative of  $\mathcal{V}_2$  computed along the solution of (1) is

$$\begin{aligned}
 d\mathcal{V}_2/dt &= (1 - q)\mu - \mu(S - S^+) - \mu(1 - q)S^+/S + \sum_{j=1}^2 \left( z_{\tau_j} e_{\tau_j} E_{\tau_j}^+ + f_{\tau_j} w_{\tau_j} I_{\tau_j}^+ + g_{\tau_j} a_{\tau_j} A_{\tau_j}^+ \right) + \sum_{k=1}^n \varphi_k \mu (q_k + V_k^+) \\
 &+ \sum_{j=1}^n (z_j - 1)(\beta_j I_j + \gamma_j A_j) S - \sum_{k=1}^{n-1} \sum_{j=k+1}^n (\varphi_k - z_k)(\beta_j I_j + \gamma_j A_j) V_k - (1 - p) \sum_{j=1}^2 f_{\tau_j} \lambda_{\tau_j} I_{\tau_j}^+ E_{\tau_j} / I_{\tau_j} \\
 &- p \sum_{j=1}^2 g_{\tau_j} \lambda_{\tau_j} A_{\tau_j}^+ E_{\tau_j} / A_{\tau_j} - \sum_{j=1}^2 z_{\tau_j} E_{\tau_j}^+ \left( (\beta_{\tau_j} I_{\tau_j} / E_{\tau_j} + \gamma_{\tau_j} A_{\tau_j} / E_{\tau_j}) S + \sum_{k=1}^{\tau_j-1} V_k (\beta_{\tau_j} I_{\tau_j} / E_{\tau_j} + \gamma_{\tau_j} A_{\tau_j} / E_{\tau_j}) \right) \\
 &- \mu \sum_{k=1}^n \varphi_k (V_k + q_k V_k^+ / V_k) + \sum_{k=1}^n \left( S^+ \beta_k + \sum_{j=1}^{k-1} \varphi_j V_j^+ \beta_k - f_k w_k \right) I_k + \sum_{k=1}^n \left( S^+ \gamma_k + \sum_{j=1}^{k-1} \varphi_j V_j^+ \gamma_k - g_k a_k \right) A_k \\
 &+ \sum_{k=1}^n ((1 - p)f_k \lambda_k + pg_k \lambda_k - z_k e_k) E_k
 \end{aligned}$$

It follows from (34) and Remark 6 that  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1$  and

$$\begin{aligned}
 S^+ \beta_k + \sum_{j=1}^{k-1} \varphi_j V_j^+ \beta_k - f_k w_k &= 0, \\
 S^+ \gamma_k + \sum_{j=1}^{k-1} \varphi_j V_j^+ \gamma_k - g_k a_k &= 0, \\
 (1 - p)f_k \lambda_k + pg_k \lambda_k - z_k e_k &= \frac{1}{a_k w_k} \left( S^+ c_k \lambda_k + \sum_{j=1}^{k-1} (\varphi_j V_j^+ c_k \lambda_k) - a_k e_k w_k z_k \right) \\
 &\begin{cases} = 0, & \text{if } k \in \mathcal{S}_2, \text{ using (22),} \\ \leq e_k (\mathcal{R}_k - 1) \leq 0, & \text{if } \mathcal{R}_k \leq 1, \quad k \notin \mathcal{S}_2, \end{cases}
 \end{aligned}$$

Let  $s = S/S^+$ ,  $v_k = V_k/V_k^+$ ,  $\sum_{\tau_i} = E_{\tau_i}/E_{\tau_i}^+$ ,  $\tau_{\tau_i} = I_{\tau_i}/I_{\tau_i}^+$ ,  $\Delta_{\tau_i} = A_{\tau_i}/A_{\tau_i}^+$ . Define

$$\begin{aligned}
 \eta &= \mu S^+, \\
 u_{\tau_j} &= z_{\tau_j} \beta_{\tau_j} I_{\tau_j}^+ S^+, \\
 b_{\tau_j} &= z_{\tau_j} \gamma_{\tau_j} A_{\tau_j}^+ S^+, \\
 \sigma_{k,\tau_j} &= z_{\tau_j} \beta_{\tau_j} I_{\tau_j}^+ V_k^+, \quad \text{for } 1 \leq k \leq \tau_j - 1, \\
 \varphi_{k,\tau_j} &= z_{\tau_j} \gamma_{\tau_j} A_{\tau_j}^+ V_k^+, \quad \text{for } 1 \leq k \leq \tau_j - 1, \\
 \Upsilon_k &= \mu \varphi_k V_k^+, \quad \text{for } 1 \leq k \leq n.
 \end{aligned}$$

We have

$$\begin{aligned}
 d\mathcal{V}_2/dt \leq & -\eta\left(s + \frac{1}{s} - 2\right) - \sum_{j=1}^2 u_{\tau_j} \left(\frac{s\tau_{\tau_j}}{\Sigma_{\tau_j}} + \frac{\Sigma_{\tau_j}}{\tau_{\tau_j}} + \frac{1}{s} - 3\right) - \sum_{j=1}^2 b_{\tau_j} \left(\frac{s\Delta_{\tau_j}}{\Sigma_{\tau_j}} + \frac{\Sigma_{\tau_j}}{\Delta_{\tau_j}} + \frac{1}{s} - 3\right) \\
 & - \sum_{j=1}^2 \sum_{k=1}^{\tau_j-1} \sigma_{k,\tau_j} \left(\frac{v_k \tau_{\tau_j}}{\Sigma_{\tau_j}} + \frac{\Sigma_{\tau_j}}{\tau_{\tau_j}} + \frac{1}{v_k} - 3\right) - \sum_{j=1}^2 \sum_{k=1}^{\tau_j-1} \varphi_{k,\tau_j} \left(\frac{v_k \Delta_{\tau_j}}{\Sigma_{\tau_j}} + \frac{\Sigma_{\tau_j}}{\Delta_{\tau_j}} + \frac{1}{v_k} - 3\right) - \sum_{k=1}^n \Upsilon_k \left(v_k + \frac{1}{v_k} - 2\right),
 \end{aligned}$$

where

$$\begin{aligned}
 2\left(\eta + \sum_{k=1}^n \Upsilon_k\right) + 3\sum_{j=1}^2 \left(u_{\tau_j} + b_{\tau_j} + \sum_{k=1}^{\tau_j-1} (\sigma_{k,\tau_j} + \varphi_{k,\tau_j})\right) = & (1 - q)\mu + \mu S^+ + \sum_{j=1}^2 z_{\tau_j} e_{\tau_j} E_{\tau_j}^+ + \mu \sum_{j=1}^n \varphi_k (q_k + V_k^+) \\
 & + \sum_{j=1}^2 (f_{\tau_j} w_{\tau_j} I_{\tau_j}^+ + g_{\tau_j} a_{\tau_j} A_{\tau_j}^+)
 \end{aligned}$$

using (22). It follows that  $d\mathcal{V}_2/dt \leq 0$ . Equality holds if  $S = S^+, I_k = E_k = R_k = 0$  for all  $k \notin \mathcal{S}_2, V_k = V_k^+$  for all  $k = 1, 2, \dots, n, I_{\tau_j}/I_{\tau_j}^+ = E_{\tau_j}/E_{\tau_j}^+ = R_{\tau_j}/R_{\tau_j}^+ = 1$ . Since  $\mathcal{E}_{\mathcal{S}_2}$  is the largest invariant set in the subset of  $\mathcal{T}$  where  $d\mathcal{V}_2/dt = 0$ , its global stability follows by the LaSalle’s Invariance Principle [56].

### 3.5 Stability analysis of the equilibrium point $\mathcal{E}_{\mathcal{S}_r}$

**Theorem 10.** For  $r = 3, 4, \dots, n$ , the equilibrium point  $\mathcal{E}_{\mathcal{S}_r}$  is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_k \leq 1$  for all  $k \notin \mathcal{S}_r$  and condition (24) is satisfied.

*Proof.* The proof of Theorem 10 is similar to that of Theorem 9 by extending  $\mathcal{S}_2$  to  $\mathcal{S}_r, r = 3, 4, \dots, n$ .

**Remark 12.** Theorem 10 can be extended to a case where  $\mathcal{S}_r = \mathcal{S}_n$ .

## 4 Model for re-infected recovered and vaccinated population

There have been confirmed cases of the COVID-19 reinfections around the world [38, 40, 42]. In this section, in addition to assuming that individuals vaccinated against strain  $k$  can get infected with emerging strains  $j > k$ , we also discuss the case where individuals who recovered from strain  $k$  can be infected with emerging strains  $j > k$ . For this additional assumption, we

extend (1) to the form

$$\begin{aligned}
 dS &= \left( (1 - q)\mu - S \sum_{j=1}^n (\beta_j I_j + \gamma_j A_j) - \mu S \right) dt, \quad S(t_0) = S_0, \\
 dV_k &= \left( q_k \mu - V_k \sum_{j=k+1}^n (\beta_j I_j + \gamma_j A_j) - \mu V_k \right) dt, \quad V_k(t_0) = V_{k0}, \quad k = 1, 2, \dots, n - 1 \\
 dV_n &= \left( q_n \mu - \mu V_n \right) dt, \quad V_n(t_0) = V_{n0}, \\
 dE_1 &= (S(\beta_1 I_1 + \gamma_1 A_1) - (\mu + \lambda_1) E_1) dt, \quad E_1(t_0) = E_{10}, \\
 dE_k &= \left( S(\beta_k I_k + \gamma_k A_k) + \sum_{j=1}^{k-1} V_j (\beta_k I_k + \gamma_k A_k) + \sum_{j=1}^{k-1} R_j (\beta_k I_k + \gamma_k A_k) - (\mu + \lambda_k) E_k \right) dt, \quad E_k(t_0) = E_{k0}, \quad k = 2, 3, \dots, n \\
 dA_k &= (p \lambda_k E_k - (\mu + r_k) A_k) dt, \quad k = 1, 2, \dots, n \quad A_k(t_0) = A_{k0}, \quad k = 1, 2, \dots, n \\
 dI_k &= ((1 - p) \lambda_k E_k - (\mu + \theta_k) I_k) dt, \quad I_k(t_0) = I_{k0}, \quad k = 1, 2, \dots, n, \\
 dR_k &= \left( \theta_k I_k + r_k A_k - R_k \sum_{j=k+1}^n (\beta_j I_j + \gamma_j A_j) - \mu R_k \right) dt, \quad R_k(t_0) = R_{k0}, \quad k = 1, 2, \dots, n - 1 \\
 dR_n &= (\theta_n I_n + r_n A_n - \mu R_n) dt, \quad R_n(t_0) = R_{n0},
 \end{aligned} \tag{35}$$

The schematic diagram of model (35) is given in Fig 2.

### 4.1 Validity of the epidemic model (35)

In this section, we discuss the validity of the proposed epidemic model (35).

**Theorem 11.** *If  $S_0 > 0, V_{k0} > 0, E_{k0} > 0, A_{k0} > 0, I_{k0} > 0, R_{k0} > 0$ , then there exist a positive unique solution of (35) in the feasible region  $\mathcal{T}$  for all  $t \geq 0$ .*

*Proof.* The proof is similar to Theorem 1.

### 4.2 Existence of equilibrium points for model (35)

The disease-free equilibrium of the system (35) is the same as that of the system (1), and given by

$$\mathcal{E}_0 = \{S^0 = 1 - q, V_1^0 = q_1, \dots, V_n^0 = q_n, E_1^0 = 0, \dots, E_n^0 = 0, A_1^0 = 0, \dots, A_n^0 = 0, I_1^0 = 0, \dots, I_n^0 = 0, R_1^0 = 0, \dots, R_n^0 = 0\}. \tag{36}$$

### 4.3 Reproduction number for model (35)

It can also be shown in a similar manner that model (35) has the same reproduction numbers  $\mathcal{R}_k$  and  $\mathcal{R}_0$  as that of model (1). This shows that the number of infection in a population where every individual who recovered from strain  $k$  is immune to all possible strains is the same for the population where individuals who recovered from strain  $k$  can still be infected with emerging strains  $j > k$ . Hence, for strain  $k$  infected individual, the expected number,  $\mathcal{R}_k$  of new infections produced by individual with strain  $k$  in a susceptible population satisfying (35) is the

same as (11). Likewise, the reproduction number  $\mathcal{R}_0$  for the system (35) is obtained in a similar manner to (14).

### 4.4 Existence of endemic equilibrium points for model (35)

**Theorem 12.** *The strain m unique equilibrium point  $\mathcal{E}_m$  for the epidemic model (35) exists in the feasible region  $\mathcal{T}$  provided  $\mathcal{R}_m > 1$ . The value of  $\mathcal{E}_m$  is the same for models (1) and (35) and obtained in Theorem (2).*

**Remark 13.** *Theorem 12 shows that the existence of a single strain endemic (strain m in this case) does not depend on whether the recovered class is immune to the strain or not. Regardless of the immunity status of the recovered strain k-class to the strain, the equilibrium point of the system will be  $\mathcal{E}_m$  if  $\mathcal{R}_m > 1$ .*

Define

$$\bar{\mathcal{R}}_{\tau_k} = (1 - q + Q_k) \frac{\bar{c}_{\tau_k} \lambda_{\tau_k}}{a_{\tau_k} e_{\tau_k} w_{\tau_k}}, \quad k = 1, 2, \dots, n. \tag{37}$$

It can be shown that  $\bar{\mathcal{R}}_{\tau_k} < 1 - q + Q_k$ . Conditions for existence of other equilibrium points are given in Theorem 13.

**Theorem 13.** *The epidemic model (35) has an equilibrium point  $\mathcal{E}_{S_2}$  satisfying*

$$\begin{aligned} E_{\tau_1}^+ &= \frac{\mu}{e_{\tau_1}} \frac{\mathcal{R}_{\tau_1} - \mathcal{R}_{\tau_2}}{\frac{\mathcal{R}_{\tau_1}}{1 - q + Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}} \left(1 - \frac{\bar{\mathcal{R}}_{\tau_1}}{1 - q + Q_{\tau_1}}\right)}, \\ E_{\tau_2}^+ &= \frac{\mu}{e_{\tau_2}} \frac{\left(1 - q + Q_{\tau_2}\right) \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}} + (Q_{\tau_2} - Q_{\tau_1}) \frac{\mathcal{R}_{\tau_1}}{1 - q + Q_{\tau_1}} \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}} - \frac{\mathcal{R}_{\tau_1}}{1 - q + Q_{\tau_1}} + \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}} \frac{\bar{\mathcal{R}}_{\tau_1}}{1 - q + Q_{\tau_1}} (\mathcal{R}_{\tau_1} - 1)}{\frac{\mathcal{R}_{\tau_1}}{1 - q + Q_{\tau_1}} - \frac{\mathcal{R}_{\tau_2}}{1 - q + Q_{\tau_2}} \left(1 - \frac{\bar{\mathcal{R}}_{\tau_1}}{1 - q + Q_{\tau_1}}\right)}, \\ S^+ &= \frac{1 - q}{\mathcal{R}_{\tau_1}}, \\ V_k^+ &= \begin{cases} \frac{q_k}{\mathcal{R}_{\tau_1}}, & \text{if } k \leq \tau_1 - 1 \\ \frac{q_k \mu}{\mu + \frac{c_{\tau_2} \lambda_{\tau_2}}{a_{\tau_2} w_{\tau_2}} E_{\tau_2}^+}, & \text{if } \tau_1 \leq k < \tau_2, \\ q_k, & \text{if } k \geq \tau_2, \end{cases} \\ A_{\tau_k}^+ &= \frac{p \lambda_{\tau_k}}{a_{\tau_k}} E_{\tau_k}^+, \quad k = 1, 2, \\ I_{\tau_k}^+ &= \frac{(1 - p) \lambda_{\tau_k}}{w_{\tau_k}} E_{\tau_k}^+, \quad k = 1, 2 \\ R_{\tau_1}^+ &= \frac{\bar{c}_{\tau_1} \lambda_{\tau_1}}{a_{\tau_1} w_{\tau_1}} \frac{E_{\tau_1}^+}{\mu + \frac{c_{\tau_2} \lambda_{\tau_2}}{a_{\tau_2} w_{\tau_2}} E_{\tau_2}^+}, \\ R_{\tau_2}^+ &= \frac{\bar{c}_{\tau_2} \lambda_{\tau_2}}{\mu a_{\tau_2} w_{\tau_2}} E_{\tau_2}^+, \end{aligned} \tag{38}$$

$E_k^+ = A_k^+ = I_k^+ = R_k^+ = 0$  for  $k \neq \tau_1, \tau_2$ , in the feasible region  $\mathcal{T}$  provided

$$\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + \frac{Q_{\tau_2} - Q_{\tau_1} + \overline{\mathcal{R}}_{\tau_1}}{1 - q + Q_{\tau_1} - \overline{\mathcal{R}}_{\tau_1}} \mathcal{R}_{\tau_1}} \tag{39}$$

is satisfied.

*Proof.* The proof is similar to that of Theorem 4.

**Remark 14** Unlike Remark 13, we see here that the existence of two endemic equilibrium points depends on the immunity status of the recovered class. The exposed and infectious equilibrium points  $E_{\tau_k}^+, A_{\tau_k}^+$ , and  $I_{\tau_k}^+$  for model (35) are greater compared to that of model (1). This is because unlike model (1), model (35) suggests that the recovered population  $R_k$  is not fully immune to emerging strains  $j > k$ , increasing the possibility of populating the exposed and infected classes. In addition, condition (39) implies that  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1$ . The condition is equivalent to  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > \frac{b\mathcal{R}_{\tau_1}}{a+(b-a)\mathcal{R}_{\tau_1}}$ , where  $a = 1 - q + Q_{\tau_1} - \overline{\mathcal{R}}_{\tau_1}$  and  $b = 1 - q + Q_{\tau_2}$ .

### 4.5 Global analysis of equilibrium points for model (35)

**Theorem 14** The disease-free equilibrium  $\mathcal{E}_0$  for model (35) is locally asymptotically stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_0 < 1$  and globally stable in the feasible region if  $\mathcal{R}_0 \leq 1$ .

*Proof.* The proof is similar to that given in Theorems 6 and 7.

**Theorem 15.** The strain  $m$  equilibrium  $\mathcal{E}_m$  for the epidemic model (35) is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_k \leq 1$  for all  $1 \leq k \neq m \leq n$  and  $\mathcal{R}_m > 1$ .

*Proof.* The proof is similar to the proof of Theorem 8.

**Theorem 16.** The equilibrium point  $\mathcal{E}_{S_2}$  for model (35) is globally stable in the feasible region  $\mathcal{T}$  if  $\mathcal{R}_k < 1$  for  $1 \leq k \neq \tau_1, \tau_2 \leq n$  and condition (39) is satisfied.

*Proof.* The proof is similar to the proof of Theorem 9.

## 5 Results: Covid-19 data analysis

We apply models (1) and (35) to analyze the United States daily COVID-19 cases (number of infection cases, recovery cases, and vaccination). The daily COVID-19 cases data are available on the CDC website [58] for the COVID-19 periods 04/01/2020 till present. We analyze the confirmed COVID-19 infection cases, and vaccination cases as reported by U.S. states, U.S. territories, New York City, and the District of Columbia from the previous day. Since the two recent variant of concerns (VOC) in the United States are the Delta and Omicron variants, we consider the case where  $n = 2$  using model (35), with the Delta variant as Strain  $\tau_1 = 1$  and the Omicron variant as strain  $\tau_2 = 2$ . Two analyses are performed in this section. The first analysis is shown in Section 5.1 to confirm the validity of the results derived in this work. Using published and estimated COVID-19 parameters, we confirm the stability results for the disease-free equilibrium, strain 1 equilibrium, strain 2 equilibrium, and the endemic equilibrium  $\mathcal{E}_{S_2}$ . In section 5.2, the real COVID-19 cases for the United States is analyzed using model (35).

### 5.1 Simulation results using published and estimated parameters

Using model (35), we confirm the existence, and stability of the disease-free equilibrium, strain 1 equilibrium, strain 2 equilibrium, and the endemic equilibrium  $\mathcal{E}_{S_2}$  for the case where strains 1 and 2 represent the Delta and Omicron variants, respectively. The CDC data for vaccination shows that about 63% of the population of the United States are fully vaccinated (either taken the two dozes of Pfizer or Moderna, or the single dose of Janssen) as at January 25, 2022. For

Table 3. Model parameter values.

Parameter	Units	Range	References
$\beta_1$	day <sup>-1</sup>	[0.1, 1.4007]	[37, 62–66]
$\beta_2$	day <sup>-1</sup>	[0.1, 1.6761]	Assumed
$\gamma_1$	day <sup>-1</sup>	[0.07, 0.9567]	[62, 63, 65, 66]
$\gamma_2$	day <sup>-1</sup>	[0.07, 0.9567]	Assumed
$q_1$	day <sup>-1</sup>	[0, 0.63]	CDC
$q_2$	day <sup>-1</sup>	[0, 0.63]	CDC
$\rho$	day <sup>-1</sup>	[0.4, 0.6]	CDC
$\lambda_1$	day <sup>-1</sup>	[1/7, 1/2]	[67]
$\lambda_2$	day <sup>-1</sup>	[1/7, 1/2]	[67]
$r_1$	day <sup>-1</sup>	[1/10.9, 1/8.83]	[37]
$r_2$	day <sup>-1</sup>	[1/12.4, 1/9.41]	[37]
$\theta_1$	day <sup>-1</sup>	[1/10.9, 1/8.83]	[37]
$\theta_2$	day <sup>-1</sup>	[1/12.4, 1/9.41]	[37]

<https://doi.org/10.1371/journal.pone.0271446.t003>

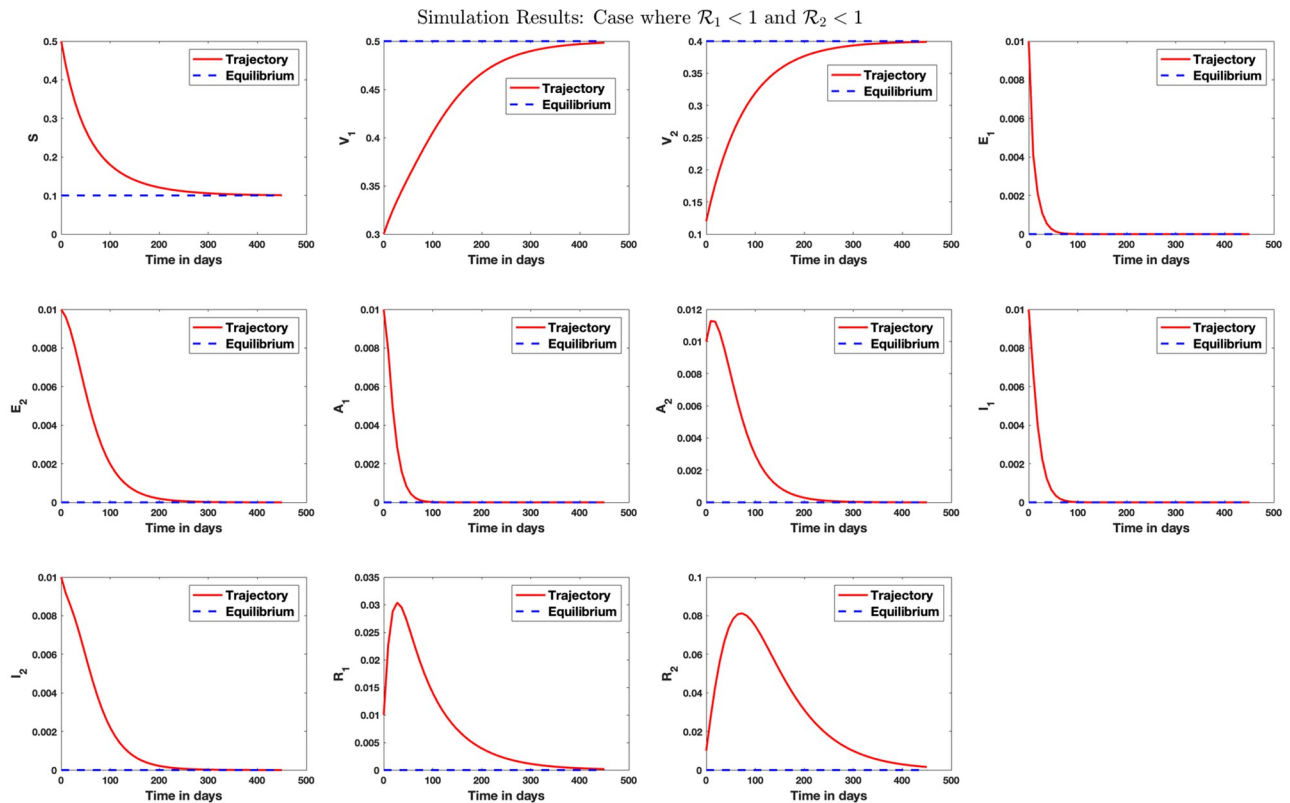
this reason, we chose  $q_1$  and  $q_2$  to be in the interval  $[0, 0.63]$ . In their paper, Saldana [59] shows that about 80% of infection was asymptomatic with average incubation and recovery time of 5 and 10 days, respectively. Also, Hay et al. [37] shows that the mean duration of Delta and Omicron's infections is 10.9 days and 9.87 days, with 95% confidence intervals (8.83, 10.9) and (9.41, 12.4), respectively. For this reason, we set  $r_1, \theta_1 \in [1/10.9, 1/8.83]$  and  $r_2, \theta_2 \in [1/12.4, 1/9.41]$ . Based on the five COVID-19 Pandemic Planning Scenarios estimated by CDC, the number of infections that are asymptomatic is uncertain and in the interval  $[0.15, 0.7]$ , with the best estimate of 30%. We set  $\mu = 0.0124$ . Bernal et al. [60] shows in their studies that the effectiveness of two doses of BNT162b2 (Pfizer) vaccines was 88.0% (95% CI, 85.3 to 90.1) among those with the delta variant. On November 16, 2020, the company Moderna announced that their vaccine is more than 94% effective at preventing COVID-19, based on an analysis of 95 cases [61]. We use these estimates for the Herd immunity plot. Following results from Jing et al, we select values for the incubation period to be between 2 days and 7 days, so that  $\lambda_1, \lambda_2 \in [1/7, 1/2]$ . The range of parameters used in the simulation is shown in Table 3.

Simulation result for the case where the population is free of disease on the long run is shown in Fig 3.

Fig 4 shows simulation result for the case where only strain 1 endemic exists in the population. Similarly, Fig 5 shows simulation result where only strain 2 endemic exists in the population. In Fig 6, we answer the question as to whether it is possible to have an endemic with more than one strain of the virus. The figure shows that this is possible if condition (39) is satisfied. Fig 7 shows the importance of condition (39) by confirming that no two strains remain in the population on the long run even if  $\mathcal{R}_1 > \mathcal{R}_2 > 1$  but  $\mathcal{R}_2 < 1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{q_2 - q_1 + \mathcal{R}_1}{1 - q_1 - \mathcal{R}_1} \mathcal{R}_1}$ . Fig 8

shows the simulation result for the case where  $\mathcal{R}_2 > \mathcal{R}_1 > 1$ .

**5.1.1 What happens if  $\mathcal{R}_1 > \mathcal{R}_2 > 1$  but condition (39) is not satisfied?** We study a case where  $\mathcal{R}_1 > \mathcal{R}_2 > 1$  but condition (39) is not satisfied. Although the condition is similar to that in Fig 6, we see here that the system converges to the strain 1 equilibrium point. That is, even though the reproduction numbers  $\mathcal{R}_1$  and  $\mathcal{R}_2$  are more than one, strain 2 still gets eradicated from the system on the long run while strain 1 caused an endemic. This study shows that the second inequality in condition (39) is necessary for the existence of endemic with more than one strain.



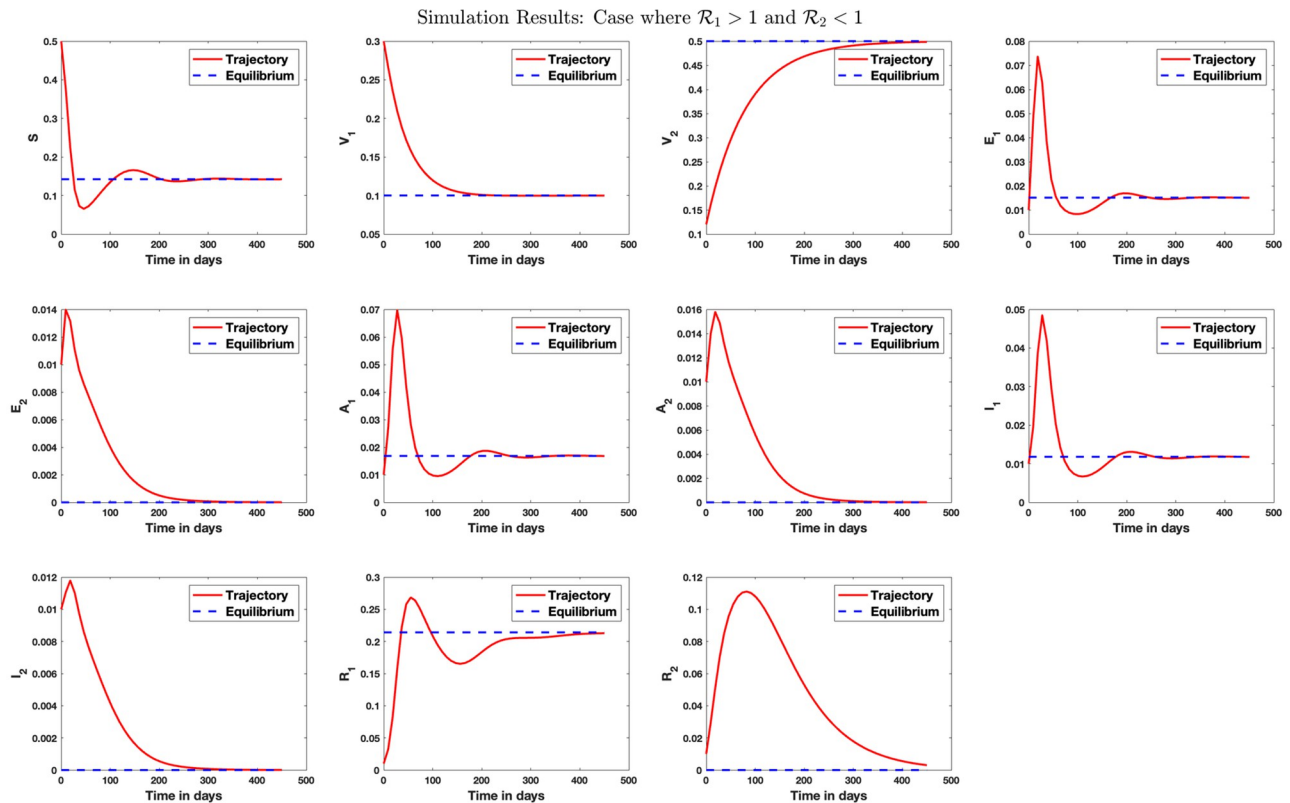
**Fig 3. Stability analysis for the disease-free equilibrium.** Case where  $\mathcal{R}_1 < 1$  and  $\mathcal{R}_2 < 1$ . Here, we set  $\beta_1 = 0.07, \beta_2 = 0.2, \gamma_1 = 0.1, \gamma_2 = 0.12, \mu = 1/80.3, q_1 = 0.5, q_2 = 0.4, \lambda_1 = 1/5, \lambda_2 = 1/4, r_1 = 1/10.5, r_2 = 1/9, \theta_1 = 0.09, \theta_2 = 0.1$ . The reproduction numbers  $\mathcal{R}_1 = 0.08$  and  $\mathcal{R}_2 = 0.74$  so that  $\mathcal{R}_0 = 0.74$ . On the long run, the susceptible population reduces to 10% of the population size while the population receiving vaccination against strains 1 and 2 increases to 50% and 40% of the population size, respectively, in this scenario. The exposed and infected population size converges to zero on the long run. The disease-free equilibrium is calculated as  $\mathcal{E}_0 = \{0.1, 0.5, 0.4, 0, 0, 0, 0, 0, 0, 0, 0, 0\}$ .

<https://doi.org/10.1371/journal.pone.0271446.g003>

**5.1.2 What happens if  $\mathcal{R}_1 > \mathcal{R}_2 > 1$ .** We study an interesting case where although the reproduction number  $\mathcal{R}_1$  for strain 1 is greater than 1, the strain still gets eradicated from the system on the long run. This happens because a newer strain 2 caused an endemic with a higher reproduction number  $\mathcal{R}_2 > \mathcal{R}_1 > 1$ . This case is analyzed geometrically in Appendix B in *S1 Appendix*.

### 5.2 Real data: Analysis using the Delta variant cases for the period 10.09.2021 to 12.18.2021

The work done here is applied to real data using the overall SARS-CoV-2 weekly Variant Proportions [68] for the United States collected from the Centers for Disease Control and Prevention (CDC) for the period 10.09.2021 to 12.18.2021. The proportion of the Delta (B.1.617.2) variant is collected and used together with the number of daily cases (collected from CDC [58]) of all variants in the United States. Within this period, data shows that on the average, the proportion of the Delta variant is about 95% of the total cases in the United State, suggesting that the Delta variant is the dominating variant during the period. For this reason, we set  $n = 1$  in (35) and estimated the parameters in Table 4 using the Nelder-Mead simplex algorithm as described in Lagarias et al. [69]. These parameters are estimated by minimizing the sum of square error between the real Delta infection cases and the simulated cases, where the



**Fig 4. Stability analysis for the strain 1 equilibrium.** Case where  $\mathcal{R}_1 > 1$  and  $\mathcal{R}_2 < 1$ . This is a case where  $\beta_1 = 1.2, \beta_2 = 0.4, \gamma_1 = 0.5, \gamma_2 = 0.08, \mu = 1/80.3, q_1 = 0.1, q_2 = 0.5, \lambda_1 = 1/5, \lambda_2 = 1/4, r_1 = 1/10.5, r_2 = 1/9, \theta_1 = 0.09, \theta_2 = 0.1$ . A scenario where the population of those receiving vaccination against strain 1 (with high transmission rate) dropped, leading to strain 1 endemic. In this case,  $\mathcal{R}_1 = 2.81$  and  $\mathcal{R}_2 = 0.86$  so that  $\mathcal{R}_0 = 2.81$ . The strain 1 equilibrium is obtained as  $\mathcal{E}_1 = \{0.1422, 0.1000, 0.5000, 0.0151, 0, 0.0168, 0, 0.0118, 0, 0.2141, 0\}$ .

<https://doi.org/10.1371/journal.pone.0271446.g004>

real Delta infection case is generated by multiplying the weekly proportion of the Delta variant by the weekly COVID-19 cases (generated from the daily cases [58]). We use  $N = 331893745$  for the total population of the United States, with initial point  $\tilde{y}_0 = (S_0, V_{10}, E_{10}, A_{10}, I_{10}, R_{10}) = (0.9076, 0.0125, 0.0001, 0.0001, 0.0019, 0.07) \times N$ , and endemic equilibrium point  $\mathcal{E}_1 = (0.5108, 0.0104, 0.2762, 0.0351, 0.1237, 0.0439) \times N$ . The real and estimated weekly COVID-19 cases for the United States for period 10.09.2021 to 12.18.2021 are given in Fig 9 and generated using model (35).

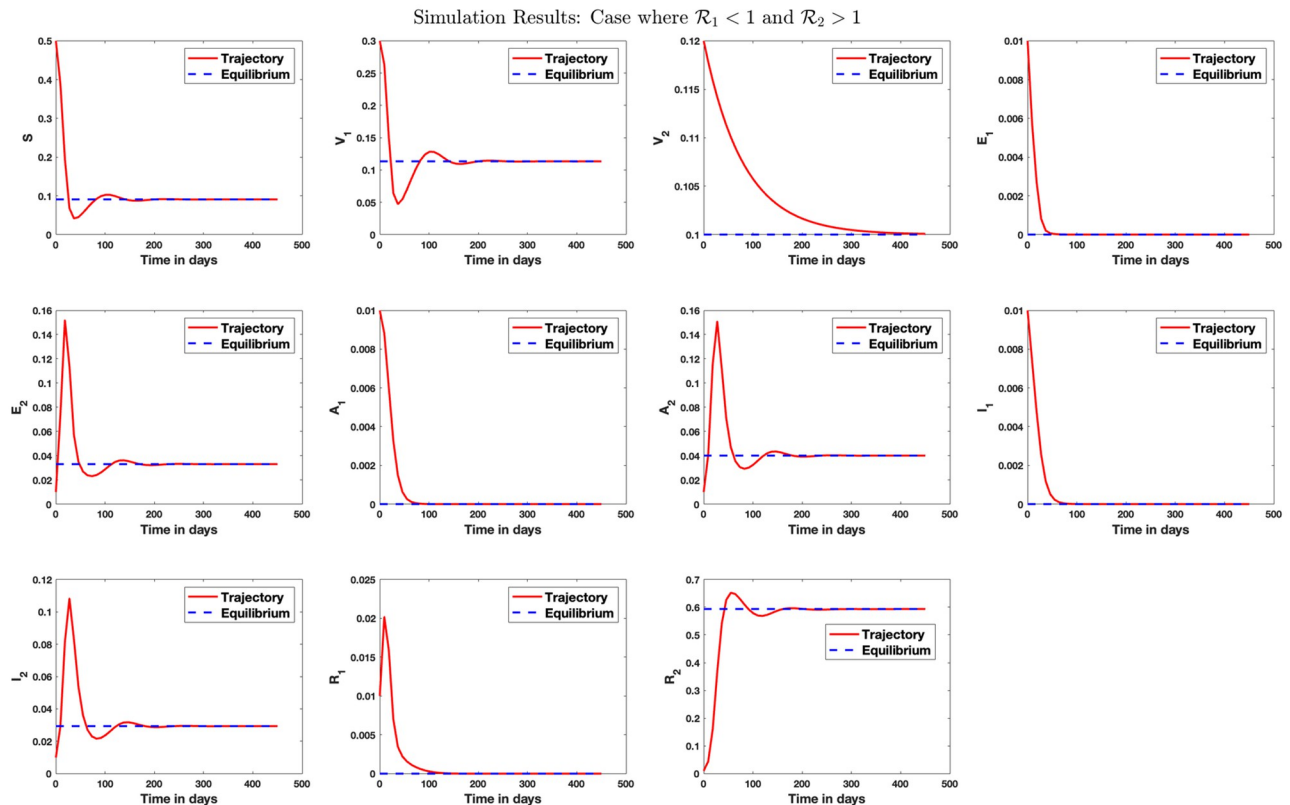
Fig 10 shows the estimated results for the population sizes  $S, V_1, E_1, A_1, I_1,$  and  $R_1$  for the Delta variant.

Using 88% and 70% for the vaccine effectiveness of two doses of Pfizer vaccine among those with Delta and Omicron variants, we calculate the Herd Immunity threshold bound to be [97%, 1). We give a plot of the Herd immunity as a function of the measure of the effectiveness of available vaccines in Fig 11.

### 5.3 Real data: Analysis using the Delta and Omicron variant cases for the period 12.11.2021 to 01.15.2022

The World Health Organization (WHO) and CDC classified a new variant, B.1.1.529, as a VOC and named it Omicron. The work done here is applied to real data using the SARS-CoV-2 weekly Variant Proportions [68] for the United States collected by CDC. The proportion of

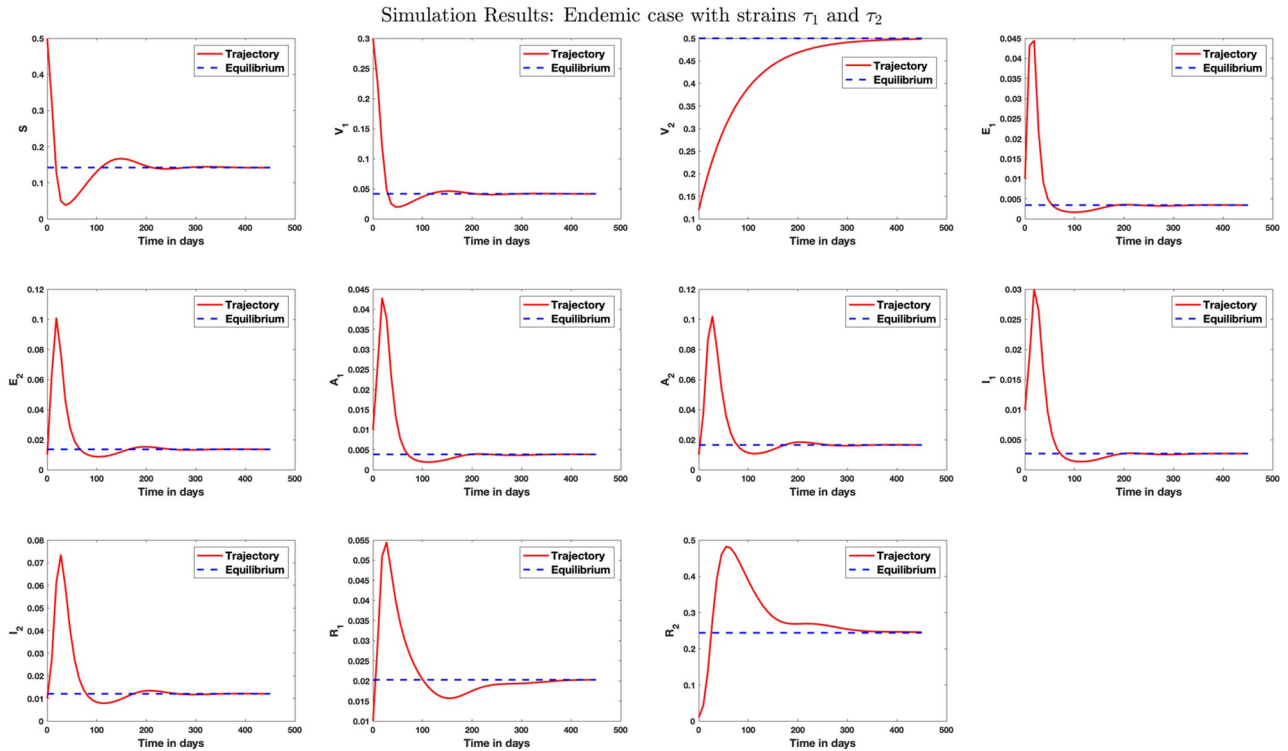




**Fig 5. Stability analysis for strain 2 equilibrium.** Case where  $\mathcal{R}_1 < 1$  and  $\mathcal{R}_2 > 1$ . In this case, we set  $\beta_1 = 0.18, \beta_2 = 0.9, \gamma_1 = 0.1, \gamma_2 = 0.4, \mu = 1/80.3, q_1 = 0.5, q_2 = 0.1, \lambda_1 = 1/5, \lambda_2 = 1/4, r_1 = 1/10.5, r_2 = 1/9, \theta_1 = 0.09, \theta_2 = 0.1$ . In this case,  $\mathcal{R}_1 = 0.47$  and  $\mathcal{R}_2 = 4.41$  so that  $\mathcal{R}_0 = 4.41$ . The strain 2 equilibrium is obtained as  $\mathcal{E}_2 = \{0.0907, 0.1134, 0.1000, 0, 0.0330, 0, 0, 0.0401, 0, 0.0294, 0, 0.5934\}$ . This vector shows proportions that each of the population sizes  $S, V_1, V_2, E_1, E_2, A_1, A_2, I_1, I_2, R_1, R_2$  converge to on the long run.

<https://doi.org/10.1371/journal.pone.0271446.g005>

the Delta (B.1.617.2) and the Omicron (B.1.1.529) variants shown in Fig 12 are collected and used together with the number of weekly cases (collected from CDC [58]) of each variants in the United States. Data suggests the existence of the two variants as VOC and that the infectivity of the Delta variant is slowing down while the infectivity of the Omicron variant is increasing during the period 12.11.2021 to 01.15.2022. For this reason, we set  $n = 2$  in (35) and estimated the parameters  $\hat{\beta}_1, \hat{\beta}_2, \hat{\gamma}_1, \hat{\gamma}_2, \hat{\mu}, \hat{q}_1, \hat{q}_2, \hat{p}, \hat{\lambda}_1, \hat{\lambda}_2, \hat{r}_1, \hat{r}_2, \hat{\theta}_1,$  and  $\hat{\theta}_2$  in Table 5 using the Nelder-Mead simplex algorithm [69]. The Delta and Omicron infected data are generated by multiplying the weekly proportion of the Delta variant by the weekly COVID-19 cases. We use  $N = 331893745$  for the total population of the United States, with initial point  $\tilde{y}_0 = (S_0, V_{10}, V_{20}, E_{10}, E_{20}, A_{10}, A_{20}, I_{10}, I_{20}, R_{10}, R_{20}) = (0.6000, 0.0187, 0.0014, 0.0010, 0.0050, 0.0010, 0.0050, 0.0023, 0.0002, 0.0070, 0.0070) \times N$ . The real and estimated weekly COVID-19 cases for the infected cases  $I_1$  and  $I_2$  for period 12.11.2021 to 01.15.2022 are given in Fig 13 and generated using model (35). The reproduction number  $\mathcal{R}_1 = 0.0676$  for the delta variant is less than one. On the other hand, the infectivity of the Omicron variant in this period was so high that the reproduction number obtained was  $\mathcal{R}_2 = 5.9503$ . The strain 2 endemic equilibrium point  $\mathcal{E}_1 = (0.1411, 0.0018, 0.1500, 0, 0.2864, 0, 0.0830, 0, 0.1274, 0, 0.2104) \times N$  suggests that the Delta variant's infectivity and exposure will die out on the long run, while there is an Omicron variant endemic. The proportion of the population vaccinated against the delta variant converges to 0.18% on the long run, and that of the Omicron variant converges to 15% on



**Fig 6. Stability analysis for the equilibrium  $\mathcal{E}_{S_2}$ .** Case where  $\mathcal{R}_1 > \mathcal{R}_2 > 1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}}$ . In this case, we set  $\beta_1 = 1.2, \beta_2 = 0.9, \gamma_1 = 0.5, \gamma_2 = 0.4, \mu = 1/80.3, q_1 = 0.1, q_2 = 0.5, \lambda_1 = 1/5, \lambda_2 = 1/4, r_1 = 1/10.5, r_2 = 1/9, \theta_1 = 0.09, \theta_2 = 0.1$ . In this case,  $\mathcal{R}_1 = 2.81$  and  $\mathcal{R}_2 = 2.45$  so that  $\mathcal{R}_0 = 2.81$ . We see in this case that  $1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}} = 1.10$  and  $\mathcal{R}_1 > \mathcal{R}_2 > 1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}}$ , implying the existence of endemic with more than one strain. We see an endemic with both strains 1 and 2 because the population is already in an endemic state with strain 1 before strain 2 caused an endemic, and the number of secondary infection caused by strain 2 is more than 1 +  $\frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}}$  but not up to that caused by strain 1. The endemic equilibrium in this case is  $\mathcal{E}_{S_2} = \{0.1422, 0.0416, 0.5000, 0.0034, 0.0136, 0.0038, 0.0165, 0.0027, 0.0121, 0.0203, 0.2439\}$ .

<https://doi.org/10.1371/journal.pone.0271446.g006>

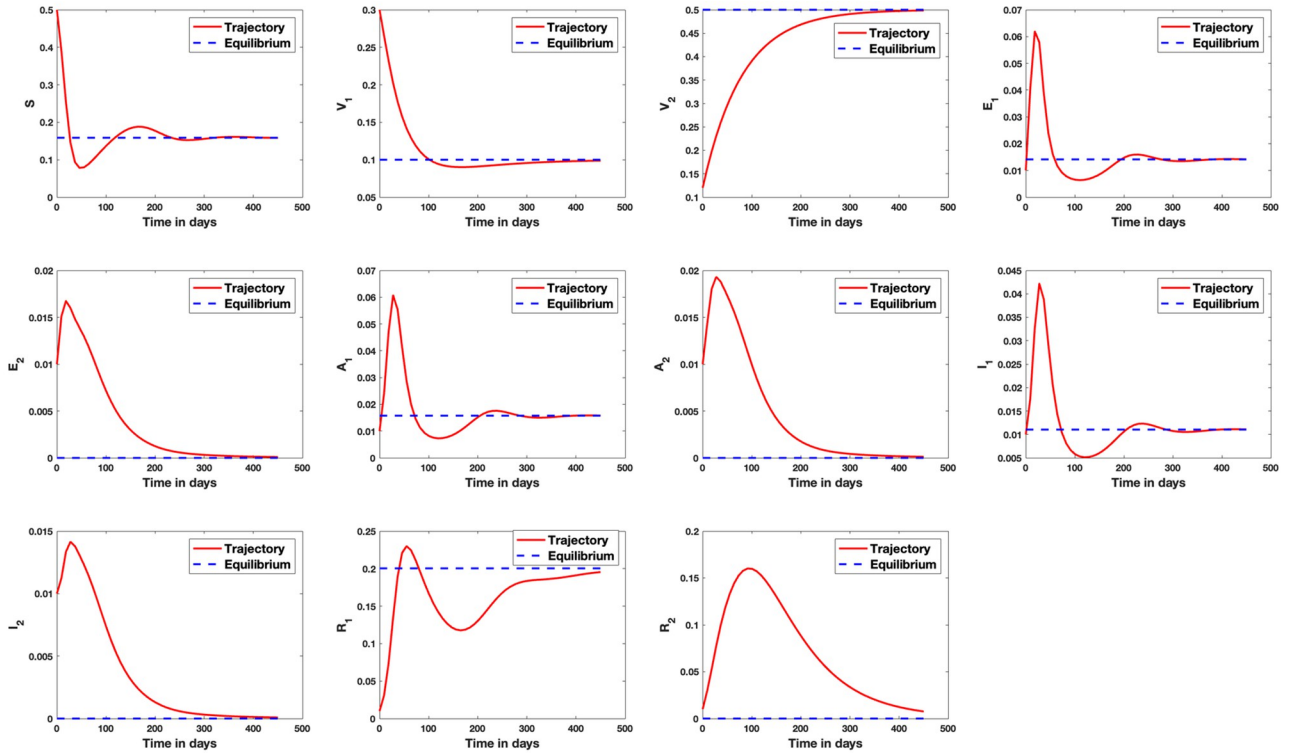
the long run. The proportion of the susceptible, exposed, asymptomatic infected, symptomatic infected, and recovered Omicron population converge to 14.1%, 28.64%, 8.30%, 12.74%, and 21.04% of the population size, respectively, on the long run.

The estimated results for all compartments are shown in Fig 14.

**5.3.1 Sensitivity analysis.** In this section, we study the impact of each epidemiological parameters on the disease transmission and prevalence. We are interested in discovering the parameters that have high impact on the basic reproduction number  $\mathcal{R}_k, k = 1, 2, \dots, n$ . This can be achieved using sensitivity analysis by calculating the sensitivity index of each parameters on  $\mathcal{R}_k$ . The normalized forward sensitivity index  $\Delta_p^u$  of a variable  $F$  that depends differentially on a parameter  $u$  is defined as [71, 72]

$$\Delta_p^u = \frac{\partial F}{\partial u} \cdot \frac{u}{F}.$$

The analytic expression for the sensitivity index of  $\mathcal{R}_k, k = 1, 2, \dots, n$ , with respect to the

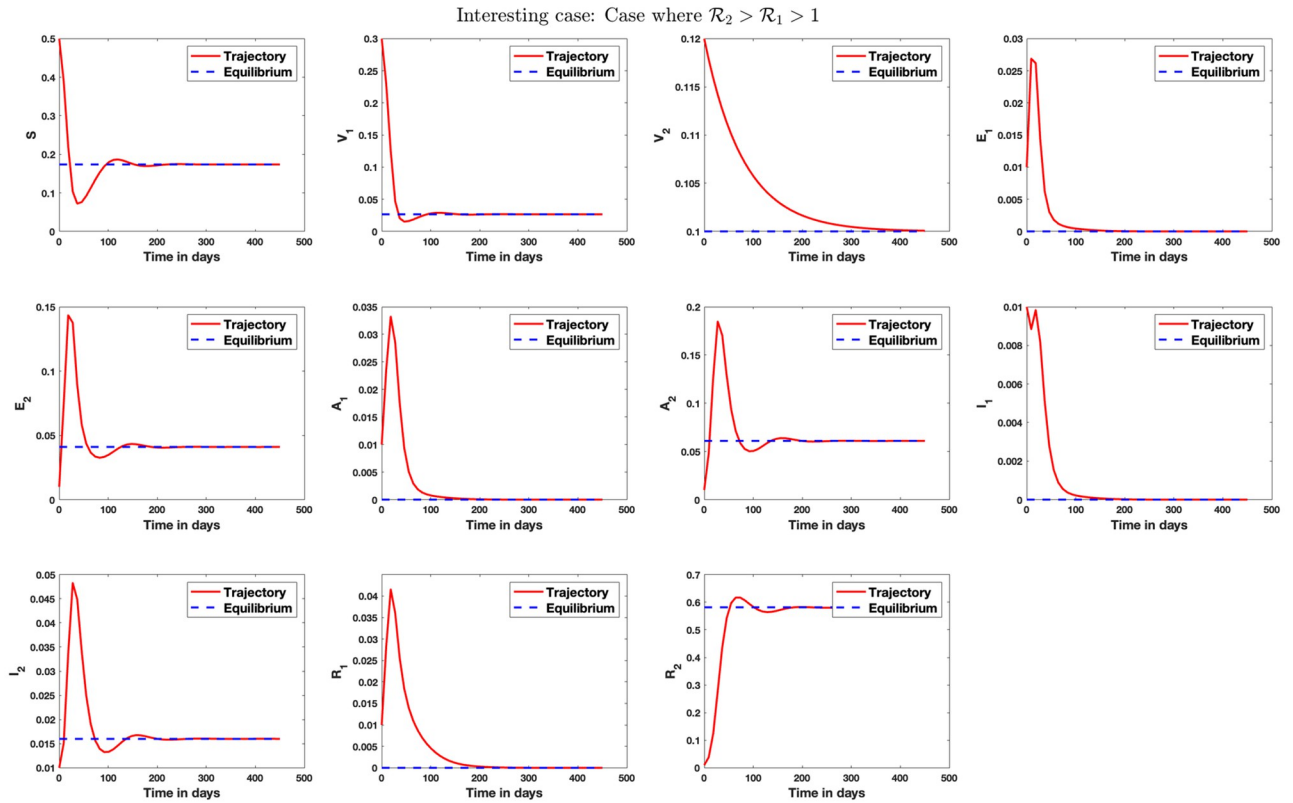


**Fig 7. Stability analysis for the case where  $\mathcal{R}_1 > \mathcal{R}_2 > 1$  but condition (39) is not satisfied.** In this case, we set  $\beta_1 = 1, \beta_2 = 0.05, \gamma_1 = 0.5, \gamma_2 = 0.4, \mu = 1/80.3, q_1 = 0.1, q_2 = 0.5, p = 0.6, \lambda_1 = 0.2, \lambda_2 = 0.25, r_1 = 1/10.5, r_2 = 1/9, \theta_1 = 0.09, \theta_2 = 0.1$ . The endemic equilibrium in this case is the strain 1 equilibrium  $\mathcal{E}_1 = \{0.1588, 0.1000, 0.5000, 0.0141, 0, 0.0158, 0, 0.0110, 0, 0.2003, 0\}$ . Here,  $\mathcal{R}_1 = 2.52, \mathcal{R}_2 = 1.01, 1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1 - \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}} = 1.09$  and condition (39) is not satisfied since  $\mathcal{R}_1 > 1 + \frac{\mathcal{R}_1 - 1}{1 + \frac{Q_2 - Q_1 + \mathcal{R}_1 - \mathcal{R}_1}{1 - q + Q_1 - \mathcal{R}_1}} > \mathcal{R}_2 > 1$ . This condition implies, from (38), that  $E_1^+ > 0$  and  $E_2^+ < 0$ , so that the values  $I_1^+, A_1^+$ , and  $R_1^+$  are positive but  $I_2^+, A_2^+$ , and  $R_2^+$  are negative. This shows that only strain 1 endemic exists on the long run. A similar analysis is presented in Appendix B in S1 Appendix geometrically.

<https://doi.org/10.1371/journal.pone.0271446.g007>

parameters in (35) is obtained as

$$\begin{aligned} \Delta_{\beta_k}^{\mathcal{R}_k} &= \frac{(1-p)a_k\beta_k}{c_k}, \\ \Delta_{\gamma_k}^{\mathcal{R}_k} &= \frac{pw_k\gamma_k}{c_k}, \\ \Delta_{\mu}^{\mathcal{R}_k} &= -\frac{1}{c_k} \left( (1-p)a_k\beta_k \left( \frac{1}{e_k} + \frac{1}{w_k} \right) + pw_k\gamma_k \left( \frac{1}{e_k} + \frac{1}{a_k} \right) \right), \\ \Delta_{q_j}^{\mathcal{R}_k} &= 0, \quad j = 1, 2, \dots, k-1, \\ \Delta_{q_{k+j}}^{\mathcal{R}_k} &= -\frac{1}{1-q+Q_k}, \quad j = 0, 1, \dots, n-k, \\ \Delta_p^{\mathcal{R}_k} &= \frac{a_k w_k}{c_k} \left( \frac{\gamma_k}{a_k} - \frac{\beta_k}{w_k} \right), \\ \Delta_{\lambda_k}^{\mathcal{R}_k} &= \frac{\mu}{\lambda_k e_k}, \\ \Delta_{r_k}^{\mathcal{R}_k} &= -\frac{pw_k\gamma_k}{a_k c_k}, \\ \Delta_{\theta_k}^{\mathcal{R}_k} &= -\frac{(1-p)a_k\beta_k}{w_k c_k}. \end{aligned}$$



**Fig 8. Stability analysis for the case where  $\mathcal{R}_2 > \mathcal{R}_1 > 1$ .** In this case, we set  $\beta_1 = 0.9, \beta_2 = 1.0, \gamma_1 = 0.4, \gamma_2 = 0.45, \mu = 1/80.3, q_1 = 0.12, q_2 = 0.1, p = 0.8, \lambda_1 = 0.21, \lambda_2 = 0.2, r_1 = 1/9, r_2 = 1/10.5, \theta_1 = 0.1, \theta_2 = 0.09$ . Here,  $\mathcal{R}_1 = 3.09$  and  $\mathcal{R}_2 = 4.49$  so that  $\mathcal{R}_0 = 4.49$ . The endemic equilibrium in this case is  $\{0.1739, 0.0267, 0.1000, 0, 0.0410, 0, 0.0609, 0, 0.0160, 0, 0.5815\}$ .

<https://doi.org/10.1371/journal.pone.0271446.g008>

The positive sensitivity index with respect to  $\beta_k, \gamma_k,$  and  $\lambda_k$  shows that an increase in the value of the strain  $k$ 's symptomatic, asymptomatic transmission rates, and the latency rate leads to an increase in the basic reproduction number,  $\mathcal{R}_k$ , as expected. Likewise, the negative sensitivity index for  $\mu, q_l, l = k, k + 1, \dots, n, r_k,$  and  $\theta_k$  shows that an increase in the natural death rate, vaccination rate, symptomatic and asymptomatic recovery rates, respectively, leads to a decrease in  $\mathcal{R}_k$ , as expected. Also, the zero value for the sensitivity index for  $q_l, l = 1, 2, \dots, k - 1$ , shows that being vaccinated against predecessor strains  $l = 1, 2, \dots, k - 1$  does not have any impact on the current and future strains  $j \geq k$  infection count. The sensitivity index  $\Delta_p^{\mathcal{R}_k}$  is positive if  $\frac{\gamma_k}{a_k} > \frac{\beta_k}{w_k}$ . That is, an increase in the fraction of infection cases that are asymptomatic will lead to an increase in the basic reproduction number,  $\mathcal{R}_k$ , if the total number of infection caused by an asymptomatic strain  $k$  infectious individual is more than that caused by a symptomatic strain  $k$  individual. The magnitude of the sensitivity of  $\mathcal{R}_k$  to changes in these parameters can be calculated by evaluating  $\Delta_u^{\mathcal{R}_k}$  for each parameter estimates  $u$  in Tables 4 and 5 for

**Table 4. Parameter estimates for model (35) with  $n = 1$ .**

Parameter	$\hat{\beta}_1$	$\hat{\gamma}_1$	$\hat{\mu}$	$\hat{q}_1$	$\hat{p}$	$\hat{\lambda}_1$	$\hat{r}_1$	$\hat{\theta}_1$
Estimate	1.3800	0.6000	0.2045	0.01037	0.3	0.1500	0.1500	0.0300

<https://doi.org/10.1371/journal.pone.0271446.t004>

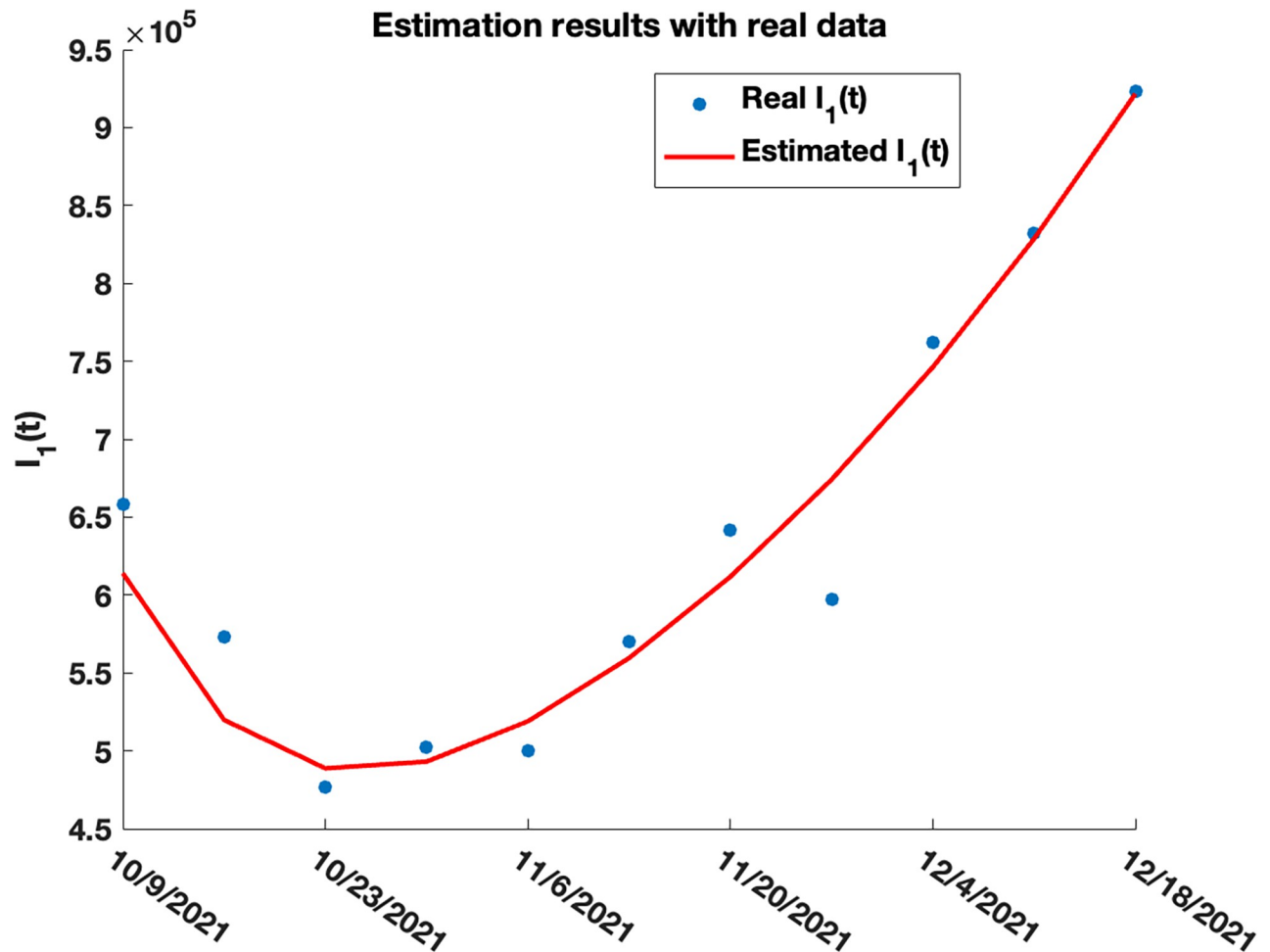


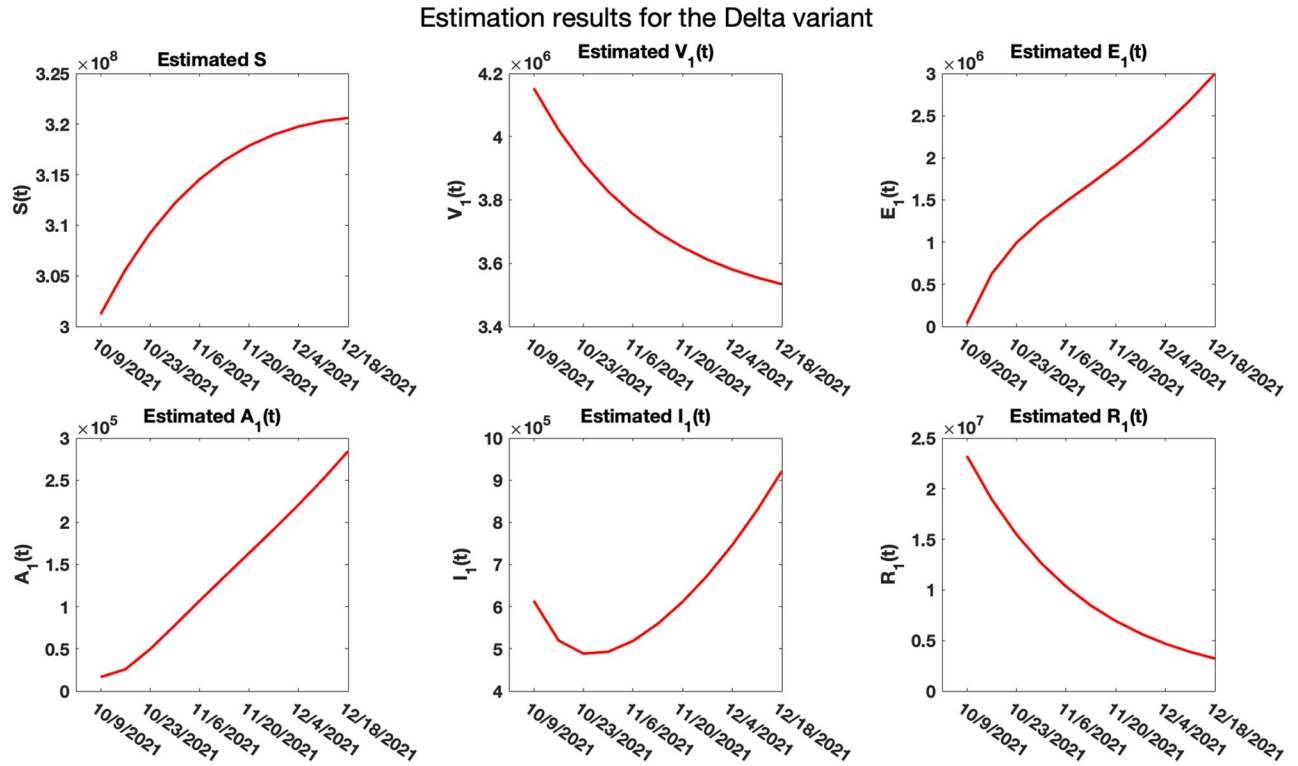
Fig 9. Real (blue) and estimated (red) COVID-19 weekly cases for the Delta variant in the United States.

<https://doi.org/10.1371/journal.pone.0271446.g009>

case where  $n = 1$  and  $n = 2$ , respectively. For a particular parameter  $u$ , a higher value of  $\Delta_u^{\mathcal{R}_k}$  suggests that  $\mathcal{R}_k$  is more sensitive to  $u$ . The sensitivity index of  $\mathcal{R}_k$  to each parameters estimated in Table 5 for model (35) is given in Table 6 for  $k = 1, 2$ .

## 6 Summary and discussion

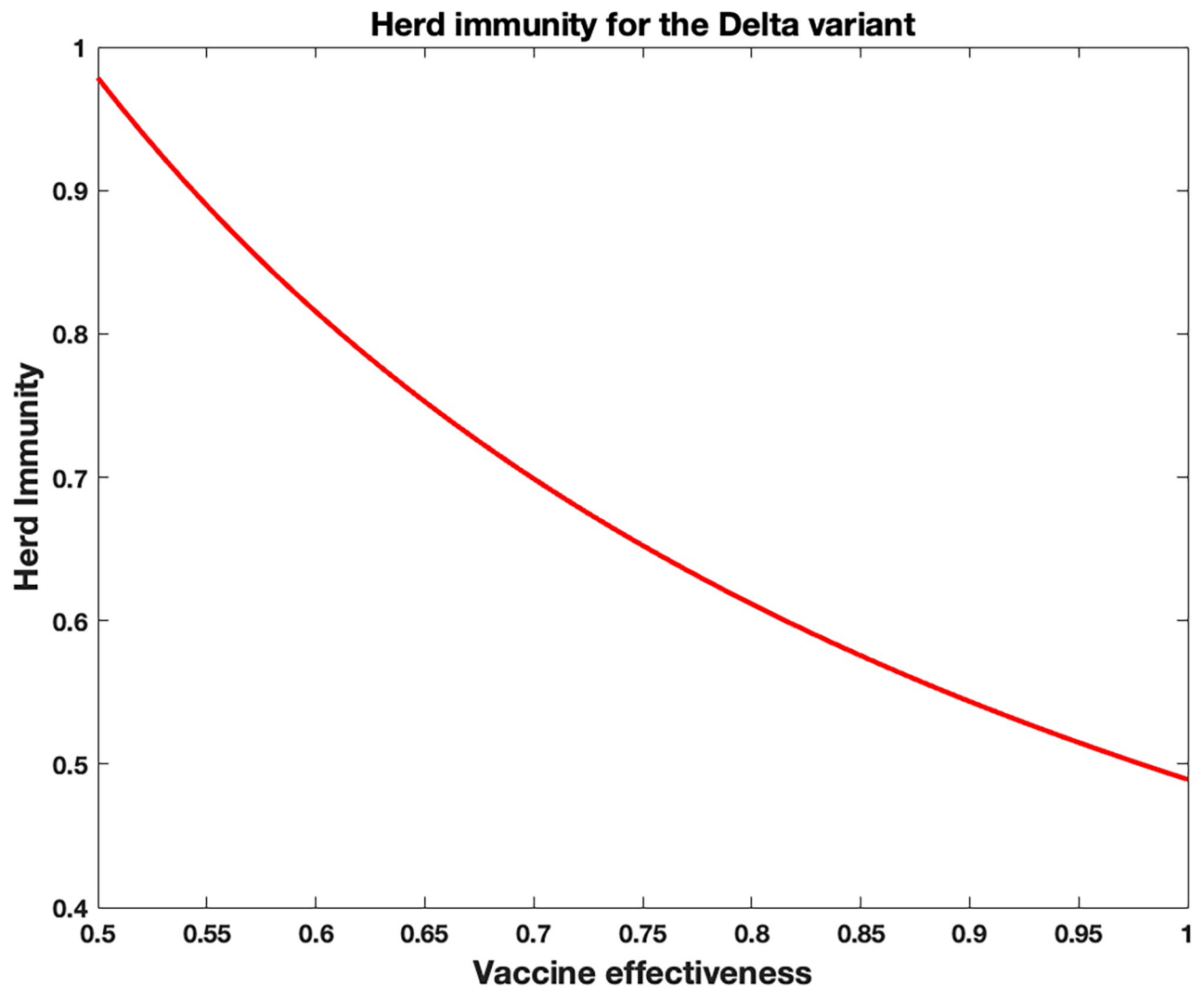
In this work, a multi-variant epidemic model for analyzing the emergence and dissemination of viral multi-strains of an infectious disease in a population that is assumed to be completely susceptible to  $n$  different strains of the disease is developed and analyzed. A major assumption on the viral strains is that those who are vaccinated and recovered from a specific strain  $k \leq n$  are immune to that strain and its predecessors but can still be infected by newer emerging strains. A study of how well-poised the model is is carried out by showing the existence of non-negative solution of the derived model. The model compares the cases where the force of infection on the susceptible vaccinated and unvaccinated populations are different and the same. The reproduction number for each specific strains  $j = 1, 2, \dots, n$  is obtained and analyzed. We show that the reproduction number for the system where individuals who recovered



**Fig 10. Estimated trajectory path for  $S, V_1, E_1, A_1, I_1,$  and  $R_1$  for the Delta variant.** The analysis suggests that the trajectories of the population of those receiving vaccination against the delta variant and those recovering from the variant are decreasing, while the susceptible, exposed and asymptomatic populations are increasing in the time period 10.09.2021 to 12.18.2021. The symptomatic population decreases from 10.09.2021 to 10.23.2021, after which it started increasing until the end of the analysis in 12.18.2021. The reproduction number  $\mathcal{R}_0$  for the variant, using (14), was calculated to be 1.94.

<https://doi.org/10.1371/journal.pone.0271446.g010>

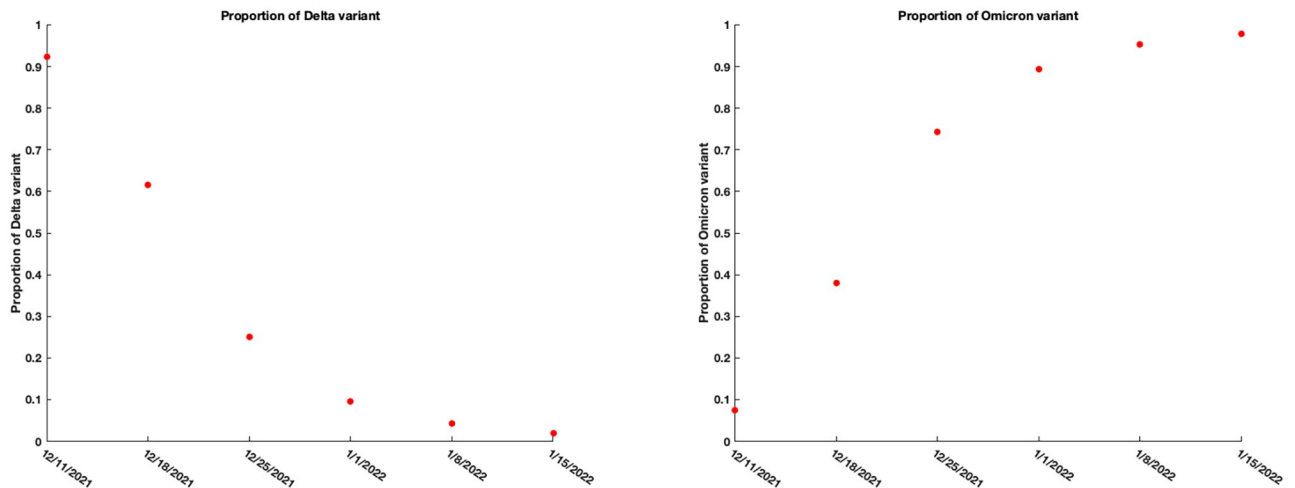
from certain strain can be infected by emerging strain is the same for the system where such individuals cannot be re-infected by emerging strain. In order to shed more light on the possibility of an endemic with more than one strain of the virus, global stability analysis is obtained for different equilibrium points  $\mathcal{E}_{S_r}, r = 0, 1, 2, \dots, n$ , of the system, with  $\mathcal{E}_{S_0}$  denoting the disease-free equilibrium. It was shown with respect to model (35) that for an endemic with strains  $\tau_1 \leq n$  and  $\tau_2 \leq n$  (where  $\tau_1 < \tau_2$ ) to occur, it must be that the reproduction number  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + \frac{Q_{\tau_2} - Q_{\tau_1} + \mathcal{R}_1}{1 - q + Q_{\tau_1} - \mathcal{R}_1} \mathcal{R}_{\tau_1}}$ . This condition shows that for an endemic with strains  $\tau_1 \leq n$  and  $\tau_2 \leq n$  to occur, the population must have been in an endemic state with the first emerged strain  $\tau_1$  and the number of secondary infection caused by strain  $\tau_1$  must be greater than that of strain  $\tau_2$ , with a necessary condition that  $\mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + \frac{Q_{\tau_2} - Q_{\tau_1} + \mathcal{R}_1}{1 - q + Q_{\tau_1} - \mathcal{R}_1} \mathcal{R}_{\tau_1}}$ . The importance of this necessary condition is emphasized numerically by showing that a system that only satisfies the condition  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1$  does not converge to the endemic equilibrium  $\mathcal{E}_{S_2}$  with strains  $\tau_1$  and  $\tau_2$ , but instead converges to strain  $\tau_1$  equilibrium  $\mathcal{E}_{\tau_1}$ . The later condition reduces to  $\mathcal{R}_{\tau_1} > \mathcal{R}_{\tau_2} > 1 + \frac{\mathcal{R}_{\tau_1} - 1}{1 + \frac{Q_{\tau_2} - Q_{\tau_1}}{1 - q + Q_{\tau_1}} \mathcal{R}_{\tau_1}}$  for the case where individuals who recovered from certain strain cannot be infected by emerging strain. A general condition is obtained for the case where endemic with strains  $\tau_1, \tau_2, \dots, \tau_r$  (with  $\tau_1 < \tau_2 < \dots < \tau_r$ ) exists in the population. Our



**Fig 11.** Herd immunity as a function of the measure of the effectiveness of vaccines for the Delta variant in the United States.

<https://doi.org/10.1371/journal.pone.0271446.g011>

analysis shows that for an endemic with strains  $\tau_1, \tau_2, \dots, \tau_r$  to exist, the population must first be in endemic with strain  $\tau_1$ , followed by  $\tau_2, \dots, \tau_r$ , so that condition (24) satisfied. Also, we showed numerically that a strain with a reproduction number greater than 1 can still die out on the long run if a newer emerging strain has a greater reproduction number. The effect of vaccines on the population is also analyzed. The herd immunity for each strain  $j = 1, 2, \dots, n$  is computed as a function of the effectiveness of vaccines against the strain. A threshold for the herd immunity for the case where there is endemic of more than one strain of the virus is also calculated. The result is applied to analyze real COVID-19 data for the Delta and Omicron variants in the United States. The reproduction numbers for the Delta and Omicron variants cases for the period 12.11.2021 to 01.15.2022 suggest that the Delta variant cases is dying down, while there is an endemic of the Omicron variant. Using model (35), possible trajectories for the susceptible, vaccinated, exposed, infected and recovered population are plotted by first estimating the parameters in the model. The parameters are estimated by minimizing the



**Fig 12. Proportion of the Delta (a) and the Omicron (b) variant for the period 12.11.2021 to 01.15.2022 in the United States collected from CDC2.**

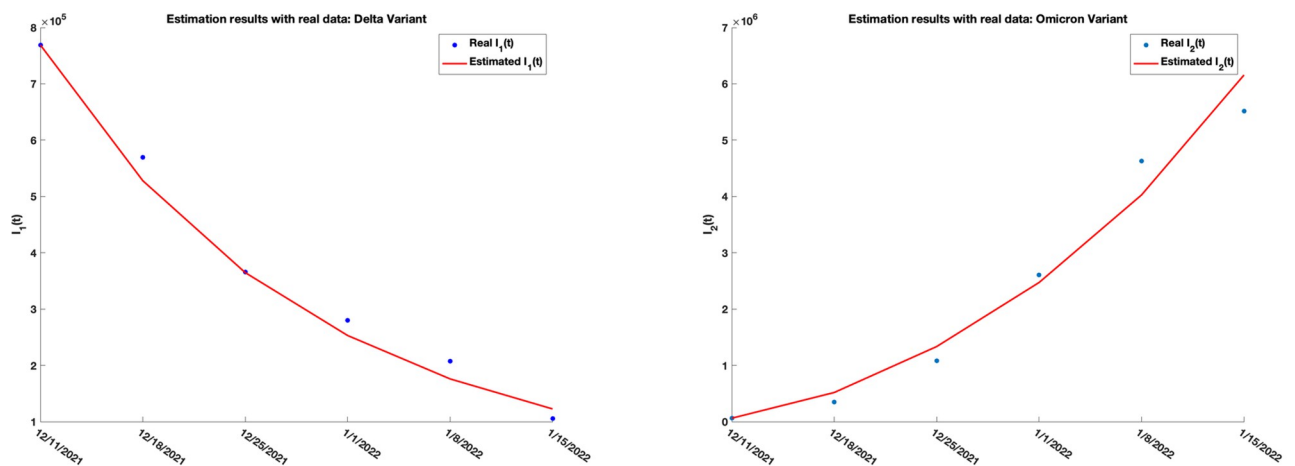
<https://doi.org/10.1371/journal.pone.0271446.g012>

**Table 5. Parameter estimates for model (35) for the Delta and Omicron variants with  $n = 2$ .**

Parameter	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\gamma}_1$	$\hat{\gamma}_2$	$\hat{\mu}$	$\hat{q}_1$	$\hat{q}_2$	$\hat{p}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{r}_1$	$\hat{r}_2$	$\hat{\theta}_1$	$\hat{\theta}_2$
Estimate	0.0397	2.1614	0.2157	5.6282	0.1500	0.0105	0.1500	0.3946	0.2550	0.2204	0.9987	0.1500	0.2972	0.1500

<https://doi.org/10.1371/journal.pone.0271446.t005>

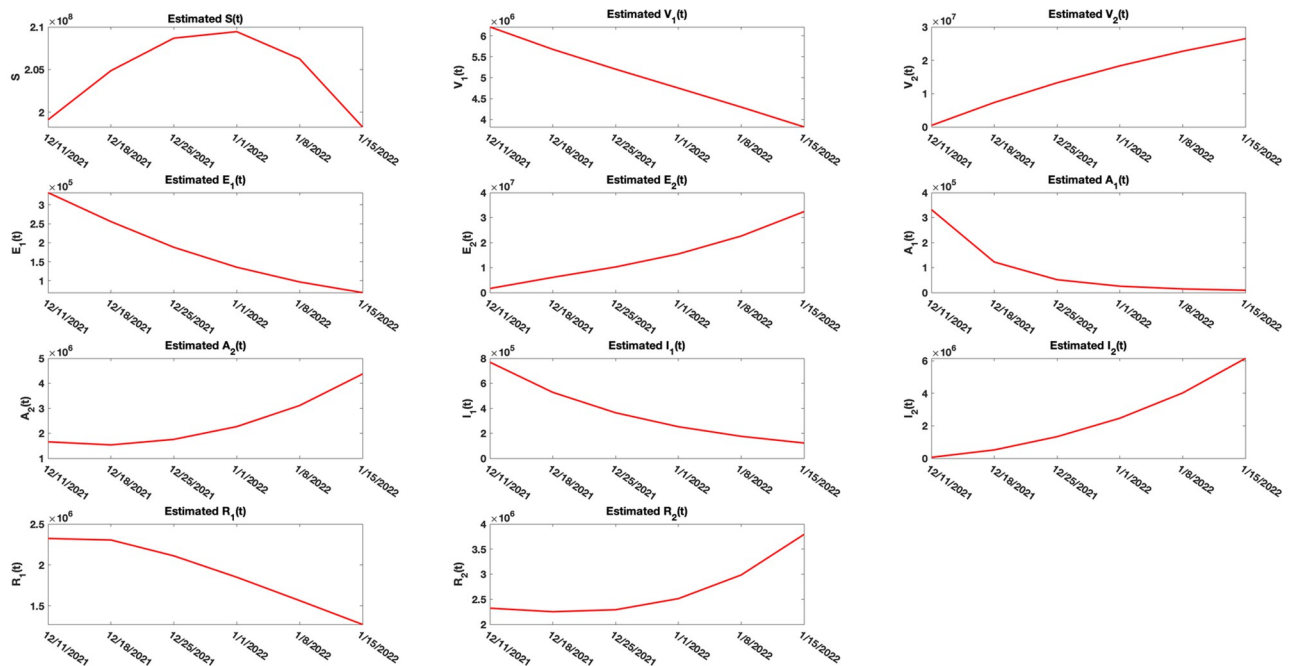
sum of square error between the real and estimated infection cases using the Nelder-Mead simplex algorithm [69]. Further research on the COVID-19 cases is ongoing and the results of the research will be made known once it is available. In the future, we plan to modify the model to fit emerging characteristics of the virus and extends limitations in the model to a more general case.



**Fig 13. Real (dotted blue) and estimated (red) COVID-19 weekly cases for the Delta variant and the Omicron variant for the period 11.27.2021 to 01.15.2022 in the United States.**

<https://doi.org/10.1371/journal.pone.0271446.g013>





**Fig 14. Estimated results for the Delta and Omicron variants.** The result shows that the number of Delta exposure and infection cases is decreasing (slowing down) while the number of Omicron exposure and infection cases is increasing (speeding up) between December 11, 2021 and January 15, 2022. These bring about a decrease in the population of those who are vaccinated against the Delta variant and an increase (at a slowing pace) in the number of those vaccinated against the Omicron variant. The number of those susceptible to the variants rises between the period 12.11.2021 to 01.01.2022, after which it started falling until 01.15.2022. The population of those who recovered from the Delta variant is speedily decreasing while the count of those who recovered from Omicron variant is increasing at a speeding pace. The result of the trajectory of the estimated Delta and Omicron cases is similar to the plot shown in the work of Nyberg et al. [70].

<https://doi.org/10.1371/journal.pone.0271446.g014>

**Table 6. Sensitivity index of  $\mathcal{R}_k$  to parameters generated in Table 5 for model (35).**

	Parameters $u$								$\theta_k$
	$\Delta_u^{\mathcal{R}_k}$	$\beta_k$	$\gamma_k$	$\mu$	$q_k$	$p_k$	$\lambda_k$	$r_k$	
$k = 1$	$\Delta_u^{\mathcal{R}_1}$	0.4204	0.5796	-3.9138	-1.1912	0.7744	1.4524	-0.5046	-0.9401
$k = 2$	$\Delta_u^{\mathcal{R}_2}$	0.3707	0.6293	-5.0312	-1.1765	0.9823	1.8374	-2.0975	-1.2358

Analysis suggests that changes in the asymptomatic transmission rate contribute more to the changes in the reproduction number  $\mathcal{R}_k$  than the symptomatic transmission rate for the Delta ( $k = 1$ ) and Omicron ( $k = 2$ ) variants. Also,  $\mathcal{R}_k$  is more sensitive to changes in the infection rate  $\lambda_k$  than the transmission rates and the fraction  $p_k$  of cases that are asymptomatic. For the Delta variant, the reproduction number  $\mathcal{R}_1$  is more sensitive to the recovery rate  $\theta_1$  of the symptomatic class than  $r_1$ . Likewise for the Omicron variant, the asymptomatic recovery rate  $r_2$  has much influence over  $\mathcal{R}_2$  than the symptomatic recovery rate  $\theta_2$ .

<https://doi.org/10.1371/journal.pone.0271446.t006>

## Supporting information

**S1 Appendix.**  
(PDF)

## Acknowledgments

I would like to acknowledge the valuable recommendations of the four anonymous reviewers who reviewed this paper. Also, special thanks to the chairman of the department of mathematics at Augusta University for his support.

## Author Contributions

**Conceptualization:** Olusegun Michael Otunuga.

**Data curation:** Olusegun Michael Otunuga.

**Formal analysis:** Olusegun Michael Otunuga.

**Methodology:** Olusegun Michael Otunuga.

**Resources:** Olusegun Michael Otunuga.

**Software:** Olusegun Michael Otunuga.

**Validation:** Olusegun Michael Otunuga.

**Writing – original draft:** Olusegun Michael Otunuga.

## References

1. US government SARS-COV-2 Interagency Group (SIG). <https://cdc.gov/coronavirus/2019-ncov/variants/variant-info.html>.
2. The United States Center for Disease and Control: Information onOMICRON variant. <https://cdc.gov/coronavirus/2019-ncov/variants/omicron-variant.html>
3. The United States Center for Disease and Control: Information on DELTA variant. <https://cdc.gov/coronavirus/2019-ncov/variants/delta-variant.html>
4. Appiah G.D., Blanton L., D'Mello T., Kniss K., Smith S., Mustaquim D., et al. Influenza activity-United States, 2014-15 season and composition of the 2015-16 influenza vaccine. *MMWR Morb Mortal Wkly Rep.* 2015; 64:21 583–590
5. Arruda E. F., Das S. S., Dias C. M., and Pastore D. H. Modelling and optimal control of multi strain epidemics, with application to COVID-19. *PLoS ONE.* 2021; 16:9. e0257512. <https://doi.org/10.1371/journal.pone.0257512> <https://doi.org/10.1371/journal.pone.0257512>
6. Bentaleb D., Harroudi S., Amine S., and Allali K. Analysis and Optimal Control of a Multistrain SEIR Epidemic Model with Saturated Incidence Rate and Treatment. *Differential Equations and Dynamical Systems* <https://doi.org/10.1007/s12591-020-00544-6>, 2020.
7. Boianelli A., Nguyen V. K., Ebensen T., Schulze K., Wilk E., Sharma N., et al. Modeling Influenza Virus Infection: A Roadmap for Influenza Research. *Viruses.* 2015; 7:5274–5304. <https://doi.org/10.3390/v7102875>
8. Diekmann O., Heesterbeek J.A.P. and Metz J.A.J. On the definition and the computation of the basic reproduction ratio  $R_0$ , in models for infectious diseases in heterogeneous populations. *J. Math. Biol.* 1990; 28:365.
9. Feng Zhilan, Towers S., and Yang Y. Modeling the Effects of Vaccination and Treatment on Pandemic Influenza. *The AAPS Journal.* 2011; 13:3 427–436. <https://doi.org/10.1208/s12248-011-9284-7> PMID: 21656080
10. Fudolig M. and Howard R. The local stability of a modified multi-strain SIR model for emerging viral strains. *PLoS ONE.* 2017; 15:12 e0243408. <https://doi.org/10.1371/journal.pone.0243408>
11. Guo D., Li K. C., Peters T. R., Snively B. M., Poehling K. A., and Zhou X. Multi-scale modeling for the transmission of influenza and the evaluation of interventions toward it. *SCIENTIFIC REPORTS.* 2015; 5:8980. <https://doi.org/10.1038/srep08980>
12. Hethcote H.W. and Driessche van den. Some epidemiological models with nonlinear incidence. *J. Math. Biol.* 1991; 29:271–287. <https://doi.org/10.1007/BF00160539>
13. Hethcote H.W. The mathematics of infectious diseases. *SIAM Review.* 2000; 42: 599–653. <https://doi.org/10.1137/S0036144500371907>
14. Kanyiri C. W., Mark K., and Luboobi L. Mathematical Analysis of Influenza A Dynamics in the Emergence of Drug Resistance. *Hindawi Computational and Mathematical Methods in Medicine.* 2018; 2018, Article ID 2434560. <https://doi.org/10.1155/2018/2434560>
15. Khyar O., and Allali K. Global dynamics of a multi-strain SEIR epidemic model with general incidence rates: application to COVID-19 pandemic. *Nonlinear Dyn.* 2020; 102:489-509 <https://doi.org/10.1007/s11071-020-05929-4> <https://doi.org/10.1007/s11071-020-05929-4>
16. Korobeinikov Andrei. Global Properties of basic virus dynamics models. *Bulletin of Mathematical Biology.* 2004; 66:4 879–883. <https://doi.org/10.1016/j.bulm.2004.02.001>

17. Korobeinikov Andrei. Lyapunov functions and global stability for SIR and SEIRS epidemiological models with non-linear transmission. *Bulletin of Mathematical Biology*. 2006; 30: 615–626.
18. Korobeinikov A. Global Properties of SIR and SEIR Epidemic Models with Multiple Parallel Infectious Stages. *Bulletin of Mathematical Biology*. 2009; 71:75–83 <https://doi.org/10.1007/s11538-008-9352-z>
19. Kucharski A.J., Llessler J., Read J.M., Zhu H., Jiang C.Q., Guan Y., et al. Estimating the life course of influenza A(H3N2) antibody responses from cross-sectional data. *PLoS Biol*. 2015. 13:3 1–16.
20. Lauer S.A., Grantz K.H., Bi Q., Jones F.K., Zheng Q., Meredith H.R., et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: Estimation and application. *Ann. Intern. Med*. 2020; 172: 577–582. <https://doi.org/10.7326/M20-0504>
21. Li J., Xiao Y., Zhang F., Yang Y. An algebraic approach to proving the global stability of a class of epidemic models. *Nonlinear Analysis: Real World Applications*. 2012; 13:2006–2016 <https://doi.org/10.1016/j.nonrwa.2011.12.022>
22. Liu J.W.M., Levin S.A. and Isawa Y. Influence of nonlinear incidence rates upon the behaviour of SIRS epidemiological models. *J. Math. Biology*. 1986; 23: 187–204. <https://doi.org/10.1007/BF00276956>
23. Liu JW. M., Hethcote H.W., and Levin S.A. Dynamical behaviour of epidemiological models with nonlinear incidence rates. *J. Math. Biology*. 1987; 25: 359–380 <https://doi.org/10.1007/BF00277162>
24. Meehan M. T., Cocks Daniel G., Trauer James M., and McBrydea Emma S. Coupled, multi-strain epidemic models of mutating pathogens. *Mathematical Biosciences*. 2018; 296:82–92. <https://doi.org/10.1016/j.mbs.2017.12.006>
25. Morris D. H., Gostic K. M., Pompei S., Bedford T., Luksza M., Neher R. A., et al. Predictive Modeling of Influenza Shows the Promise of Applied Evolutionary Biology. *Trends in Microbiology*. 2018; 26:2 <http://dx.doi.org/10.1016/j.tim.2017.09.004>
26. Otunuga O. M. Estimation of epidemiological parameters for COVID-19 cases using a stochastic SEIRS epidemic model with vital dynamics. *Results in Physics*. 2021; 28:104664. <https://doi.org/10.1016/j.rinp.2021.104664>
27. Otunuga O. M. and Ogunsolu M. O. Qualitative analysis of a stochastic SEITR epidemic model with multiple stages of infection and treatment. *Infectious Disease Modelling*. 2020; 5: 61–90 <https://doi.org/10.1016/j.idm.2019.12.003>
28. Mummert A. and Otunuga O. M. Parameter identification for a stochastic SEIRS epidemic model: case study influenza. *Journal Mathematical Biology*. 2019; 79:705–729. <https://doi.org/10.1007/s00285-019-01374-z>
29. Otunuga O. M. Global stability for a  $2n+1$  dimensional HIV/AIDS epidemic model with treatments. *Mathematical Biosciences*, 2018; 299:138–152, <https://doi.org/10.1016/j.mbs.2018.03.013> <https://doi.org/10.1016/j.mbs.2018.03.013>
30. Otunuga O. M. Global Stability of Nonlinear Stochastic SEI Epidemic Model. *International Journal of Stochastic Analysis*. 2017; 2017: 1–7. <https://doi.org/10.1155/2017/6313620>
31. Gonzalez-Parra G., Martínez-Rodríguez D., and Villanueva-Micó R. J. Impact of a New SARS-CoV-2 Variant on the Population: A Mathematical Modeling Approach. *Math. Comput. Appl*. 2021; 26, 25. <https://doi.org/10.3390/mca26020025>
32. Rahimi F., Abadi A.T.B. Implications of the Emergence of a New Variant of SARS-CoV-2, VUI-202012/01. *Arch. Med. Res*. 2021.
33. Reicha N. G., Brook L. C., Fox S. J., Kandula S., McGowan C. J., Moore E., et al. A collaborative multi-year, multimodel assessment of seasonal influenza forecasting in the United States. *PNAS*. 2019; 116:8:3146–3154 <https://doi.org/10.1073/pnas.1812594116>
34. Song H., Jia Z., Jin Z., and Liu S. Estimation of COVID-19 outbreak size in Harbin, China. *Nonlinear Dyn*. 2021; 106(2): 1229–1237 <https://doi.org/10.1007/s11071-021-06406-2> <https://doi.org/10.1007/s11071-021-06406-2>
35. Xu R., Ekiert D.C., Krause J.C., Hai R., Crowe IA. Jr, Wilson J.E. Structural basis of preexisting immunity to the 2009 H1N1 pandemic influenza virus. *Science*. 2010; 328:357–360. <https://doi.org/10.1126/science.1186430>
36. Yang Y., Sugimoto J.D., Halloran M.E., Basta N.E., Chao D.L., Matraj L., et al. The transmissibility and control of pandemic influenza A (H1N1) virus. *Science*. 2009; 326:5953 729–733 <https://doi.org/10.1126/science.1177373>
37. A. Hay J., M. Kissler S., R. Fauver J., Mack C., Tai C. G., Samant R. M., et al. 2022. Viral dynamics and duration of PCR positivity of the SARS-CoV-2 Omicron variant. Pre-print.
38. Iwasaki A. What reinfections mean for COVID-19. *Lancet Infect Dis*. 2021; 21(1)

39. Massard M., Eftimie R., Perasso A., Sausseure B. A multi-strain epidemic model for COVID-19 with infected and asymptomatic cases: Application to French data. *Journal of Theoretical Biology*. 2022; 545: 111117. <https://doi.org/10.1016/j.jtbi.2022.111117>
40. Wang J., Kaperak C., Sato T., Sakuraba A. COVID-19 reinfection: a rapid systematic review of case reports and case series. *J Investig Med*. 2021; 69:1253–1255. <https://doi.org/10.1136/jim-2021-001853>
41. Yadegari I., Omid M., and Smith S. R. The herd-immunity threshold must be updated for multi-vaccine strategies and multiple variants. *Scientific Reports* 2021; 11:22970 <https://doi.org/10.1038/s41598-021-00083-2>
42. Zhang K., Yiu-Nam Lau J., Yang L. and Zhen-Guo Ma. SARS-CoV-2 reinfection in two patients who have recovered from COVID-19. *Precision Clinical Medicine*, 3(4), 2020, 292–293 <https://doi.org/10.1093/pcmedi/pbaa031>
43. Zheng T., Nie L., Teng Z., and Luo Y. Competitive exclusion in a multi-strain malaria transmission model with incubation period. *Chaos, Solitons and Fractals*: 2020; 131, 109545 <https://doi.org/10.1016/j.chaos.2019.109545>
44. Eberl H. J., Gaebler H. J., and Grohn Y. T. A brief note on a multistrain SIR model with complete cross-protection and nonlinear force of infection. *Commun Nonlinear Sci Numer Simul* 2021; 103:106001, <https://doi.org/10.1016/j.cnsns.2021.106001>
45. Nisar K.S., Ahmad S., Ullah A., Shah K., Alrabaiah H., Arfan M. Mathematical analysis of SIRD model of COVID-19 with Caputo fractional derivative based on real data. *Results in Physics* 21 (2021) 103772 <https://doi.org/10.1016/j.rinp.2020.103772>
46. Ahmad S., Ullah A., Al-Mdallal Q. M., Khan H., Shah K., Khan A. Fractional order mathematical modeling of COVID-19 transmission. *Chaos, Solitons and Fractals*. 2020; 139: 110256 <https://doi.org/10.1016/j.chaos.2020.110256>
47. Hattaf K., Mohsen A. A., Harraq J., and Achaich Naceur. Modeling the dynamics of COVID-19 with carrier effect and environmental contamination. *International Journal of Modeling, Simulation, and Scientific Computing*. 2021; 12:3 2150048 <https://doi.org/10.1142/S1793962321500483>
48. Yousaf M., Zahir S., Riaz M., Hussain S. M., Shah K. Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan. *Chaos, Solitons and Fractals*. 2020; 138: 109926 <https://doi.org/10.1016/j.chaos.2020.109926>
49. Hattaf K. A New Generalized Definition of Fractional Derivative with Non-Singular Kernel. *Computation* 2020; 8, 49.
50. Hattaf K. Stability of Fractional Differential Equations with New Generalized Hattaf Fractional Derivative. *Hindawi Mathematical Problems in Engineering* Volume 2021, Article ID 8608447
51. Kelley W. G. and Peterson A. C. *The theory of differential equations Classical and Qualitative*. Springer New York Dordrecht Heidelberg London. 2010
52. Driessche P. V. and Watmough J. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. *Math. Biosci.* 2002; 180:29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)
53. Plemmons R. J. *M-Matrix Characterizations. I—Nonsingular M-Matrices*. *Linear Algebra and its Applications*. 1977; 18(2) 175–188. [https://doi.org/10.1016/0024-3795\(77\)90073-8](https://doi.org/10.1016/0024-3795(77)90073-8)
54. Berman A., and Plemmons R.J. *Nonnegative matrices in the mathematical sciences*. SIAM, Philadelphia, 1994.
55. Steele J. M. *The Cauchy-Schwarz Master Class: An Introduction to the Art of Mathematical Inequalities*. MAA Problem Books Series. Cambridge University Press, 2004.
56. J. P. LaSalle *The stability of dynamical systems: Regional conference series in applied mathematics*. SIAM, Philadelphia, 1976.
57. Ontario Public Health. COVID-19- What We Know So Far About Herd Immunity. Synthesis. 2021.
58. Centers for Disease Control and Prevention (CDC) Data. United States COVID-19 Cases and Deaths by State over Time. CDC 2021; <https://data.cdc.gov/Case-Surveillance/United-States-COVID-19-Cases-and-Deaths-by-State-o/9mfq-cb36>
59. Saldaña F., Velasco-Hernández J. X. Modeling the COVID-19 pandemic: a primer and overview of mathematical epidemiology. *SeMA Journal*. 2021; <https://doi.org/10.1007/s40324-021-00260-3>
60. Bernal J. L., Andrews N., Gower C., Gallagher E., Simmons R., Thelwall S., et al. Effectiveness of Covid-19 Vaccines against the B.1.617.2 (Delta) Variant. *N ENGL J MED*. 2021; 385:7.
61. Saldaña F., Velasco-Hernández J. X. The trade-off between mobility and vaccination for COVID-19 control: a metapopulation modelling approach. *R. Soc. Open Sci.*2021; 8(6): 202240. <https://doi.org/10.1098/rsos.202240>

62. Giordano G., Blanchini F., Bruno R., Colaneri P., Di Filippo A., Di Matteo A., et al. Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy. *Nature Medicine*. 2020; 26: 855–860; [www.nature.com/naturemedicine](http://www.nature.com/naturemedicine) <https://doi.org/10.1038/s41591-020-0883-7>
63. Kumari P., Singh H. P., Singh S. SEIAQRDT model for the spread of novel coronavirus (COVID-19): A case study in India. *Applied Intelligence*. 2021; 51:2818–2837. <https://doi.org/10.1007/s10489-020-01929-4> <https://doi.org/10.1007/s10489-020-01929-4>
64. Niño-Torres D., Ríos-Gutiérrez A., Arunachalam V., Ohajunwa C., Seshaiyer P. Stochastic modeling, analysis, and simulation of the COVID-19 pandemic with explicit behavioral changes in Bogotá: A case study. *Infectious Disease Modelling*. 2022; 7: 199e211
65. Shen Zhong-Hua, Chu Yu-Ming, Khan M.A., Muhammad S., Al-Hartomy O. A., Higazy M. Mathematical modeling and optimal control of the COVID-19 dynamics. *Results in Physics*. 2021; 31:105028 <https://doi.org/10.1016/j.rinp.2021.105028> PMID: 34868832
66. Sustersic T., Blagojevic A., Cvetkovic D., Cvetkovic A., Lorencin I., Segota S. B., et al. Epidemiological Predictive Modeling of COVID-19 Infection: Development, Testing, and Implementation on the Population of the Benelux Union. *Front. Public Health*, 28 October 2021; 9:727274 <https://doi.org/10.3389/fpubh.2021.727274> <https://doi.org/10.3389/fpubh.2021.727274>
67. Jing M., Ng K. W., Namee B. M., Biglarbeigi P., Brisk R. Bond R., Finlay D., et al. COVID-19 modelling by time-varying transmission rate associated with mobility trend of driving via Apple Maps. *Journal of Biomedical Informatics* 2021; 122:103905 <https://doi.org/10.1016/j.jbi.2021.103905>
68. Centers for Disease Control and Prevention (CDC) Monitoring Variant Proportions. United States COVID-19 Variant Proportions CDC 2021; <https://covid.cdc.gov/covid-data-tracker/#variant-proportions>
69. Lagarias J.C., Reeds J. A., Wright M. H., and Wright P. E. Convergence Properties of the Nelder-Mead Simplex Method in Low Dimensions. *SIAM Journal of Optimization*. 1998; 9(1): 112–147. <https://doi.org/10.1137/S1052623496303470>
70. Nyberg T., Ferguson N. M., Nash S. G., Webster H. H., Flaxman S., Andrews N., et al. Comparative analysis of the risks of hospitalisation and death associated with SARS-CoV-2 omicron (B.1.1.529) and delta (B.1.617.2) variants in England: a cohort study. *Lancet* 2022; 399:1303–12 [https://doi.org/10.1016/S0140-6736\(22\)00462-7](https://doi.org/10.1016/S0140-6736(22)00462-7)
71. Chitnisa N., Hyman J. M., Cushing J. M. Determining Important Parameters in the Spread of Malaria Through the Sensitivity Analysis of a Mathematical Model. *Bulletin of Mathematical Biology*. 2008; 70: 1272–1296. <https://doi.org/10.1007/s11538-008-9299-0>
72. Lloyd A.L. Sensitivity of Model-Based Epidemiological Parameter Estimation to Model Assumptions. *Mathematical and Statistical Estimation Approaches in Epidemiology*. 2009; PMID: PMC7121564: 123–141.