



Research article

Evaluations of heterogeneous epidemic models with exponential and non-exponential distributions for latent period: the Case of COVID-19

Huiping Zang, Shengqiang Liu* and Yi Lin

School of Mathematical Sciences, Tiangong University, Tianjin 300387, China

* **Correspondence:** Email: sqliu@tiangong.edu.cn.

Abstract: Most of heterogeneous epidemic models assume exponentially distributed sojourn times in infectious states, which may not be practical in reality and could affect the dynamics of the epidemic. This paper investigates the potential discrepancies between exponential and non-exponential distribution models in analyzing the transmission patterns of infectious diseases and evaluating control measures. Two SEIHR models with multiple subgroups based on different assumptions for latency are established: Model I assumes an exponential distribution of latency, while Model II assumes a gamma distribution. To overcome the challenges associated with the high dimensionality of GDM, we derive the basic reproduction number (R_0) of the model theoretically, and apply numerical simulations to evaluate the effect of different interventions on EDM and GDM. Our results show that considering a more realistic gamma distribution of latency can change the peak numbers of infected and the timescales of an epidemic, and GDM may underestimate the infection eradication time and overestimate the peak value compared to EDM. Additionally, the two models can produce inconsistent predictions in estimating the time to reach the peak. Our study contributes to a more accurate understanding of disease transmission patterns, which is crucial for effective disease control and prevention.

Keywords: heterogeneous epidemic models; exponential distribution; gamma distribution; latent; numerical simulations

1. Introduction

Throughout history, infectious diseases have posed a threat to human existence and even influenced the course of history [1]. For example, the Black Death epidemic in Europe in the 14th century killed more than 24 million people [2]. Now, with the outbreak of COVID-19, which has already caused unprecedented public health challenges worldwide [3], it is more important than ever to study the prevalence and spread of infectious diseases.

Mathematical modeling is a momentous approach to understand the transmission pattern of epi-

demic [4]. Establishing and analyzing epidemic models plays a significant role in revealing the epidemic law of diseases and predicting their development trends [5]. Furthermore, it also helps to identify the causes and key factors of disease epidemics and provides theoretical guidance in the search for appropriate control measures [5].

Epidemic models generally assume that the population is homogenous or randomly mixed, allowing the model to ignore meta-population heterogeneity including gender differences, regional variations, age differences, etc. [6]. However, due to the oversimplification and idealization of homogenous models, sometimes the analysis results deviate from the actual situation [4]. Especially for emerging diseases such as MERS, SARS, and COVID-19, population heterogeneity has an important impact on the epidemic law of diseases and the assessment of prevention and control measures. Therefore, it is crucial to consider heterogeneity factors in epidemiological model in order to solve practical problems [4].

In recent years, much progress has been made in investigating how heterogeneity changes or influences the dynamic behavior of infectious diseases. Lajmanovich et al. [7] described the effect of heterogeneity on infectious diseases by establishing a model of infectious diseases in n groups. On this basis, Nold [8] discussed the dynamic behavior of a class of disease models by considering the heterogeneity of infectious disease transmission in population. Wang et al. [9] considered aged heterogeneity factors of a COVID-19 model and further evaluated the effectiveness of control measures. Cui et al. [4] developed a meta-population model that considered latent and asymptomatic individuals separately and investigated the impact of heterogeneity factors. Dimarco et al. [10] analyzed the effect of heterogeneity on the transmission pattern of infectious diseases by combining social contacts and epidemic dynamics. Other studies related to heterogeneity can be found in the literatures [6, 11–16].

In spite of the above excellent literatures on heterogeneous epidemic models, however, most of these references assume exponential distribution of latency, however, for many infectious diseases, this is not realistic. Indeed, the distribution of latency maybe not exponential [17], and models with exponential distribution may lead to bias in understanding the dynamic behavior of diseases and evaluating control measures [18]. Moreover, there have been many cogent literatures supporting non-exponential latency distribution for COVID-19 [19–23]: in [23], by fitting the latency data of 109 COVID-19 cases, the results show that the gamma distribution may be the optimal distribution of COVID-19 latency. Therefore, it is meaningful to consider that latency obeys a non-exponential distribution in the study of infectious diseases.

In recent years, there have been some studies generated on latency obeying a non-exponential distribution. Feng et al. [24] developed a model for disease stages following a general distribution and investigated the impact of disease stages following exponential and non-exponential distributions on the aspects of evaluating the effectiveness of interventions. Feng et al. [25] further derived the basic reproduction number of disease stages obeying a non-exponential distribution, providing threshold conditions for outbreak control and prevention. Capistran et al. [26] presented a COVID-19 mathematical model with a non-exponential distribution of disease stages, in order to forecast hospital demand in urban regions throughout a COVID-19 pandemic and estimated lockdown-induced second fluctuations. The above-mentioned literatures provide an in-depth study of epidemic models in which the latent duration admits a non-exponential distribution but does not consider the influence of heterogeneous factors on the model.

On the other hand, Blyuss et al. [27] used an SEIR-type heterogeneous mathematical model as-

suming non-exponentially distributed disease stages, in order to explore the dynamics and control of COVID-19. Chen et al. [28] investigated the effect of mathematical models with heterogeneous susceptible populations when the infectious period is non-exponentially distributed on studying COVID-19 transmission dynamics. None of the above studies have considered the potential discrepancies between exponentially and non-exponentially distributed heterogeneous models, when applying models to reveal patterns of transmission of infectious diseases, predict their trends, and evaluate control measures. It remains a worthwhile question of how the non-exponential distribution for latency of heterogeneous models impacts the outbreak transmission dynamics and interventions compared to exponential distributions. To our knowledge, this problem has not been well researched.

In this paper, motivated by the above issues, we investigate the potential differences between exponentially distributed and non-exponentially distributed heterogeneous models when used to understand patterns of transmission of infectious diseases and to evaluate control measures. To the best of our knowledge, compared to the existing literatures, the main innovation of this paper is that it considers for the first time the differences between latent period obeying exponential and non-exponential distributions in heterogeneous models. It is believed that our research seems to bring new viewpoint for the study and control of diseases.

The paper is organized as follows: Section 2 includes the establishment of the heterogeneous model with m subpopulations whose disease stages obey a general distribution. Then, we reduce the general distribution model into exponential distribution model (EDM) and gamma distribution model (GDM). Furthermore, the basic reproduction number R_0 is given a detailed theoretical derivation. In Section 3, PRCC (Partial rank correlation coefficient) is used to examine how the sensitivity of parameters changes during disease transmission, and possible control strategies are evaluated by sensitivity analysis. Next, in order to investigate how control parameters affect the dynamic behavior of infectious diseases and conclude differences in the assessment of the models, we use numerical simulations to compare EDM and GDM. Section 4 involves some discussion and conclusion remarks.

2. Model and analysis

2.1. The model with general distribution of disease stages

Table 1. Explanation of symbols.

Symbol	Description
N_i	Total population of group i
S_i	Susceptible individuals of group i
E_i	Exposed/Latent individuals of group i
I_i	Infectious individuals of age group i
H_i	Hospitalized individuals of group i
R_i	Recovered/removed individuals of age group i
$Q_i(t)$	Probability that an individual in group i remains in the latent period for t time units since entering E_i stage
$P_i(t)$	Probability that an individual in group i remains infectious for t time units since onset
$L_i(t)$	Probability that an individual in group i has not been hospitalized for t time units since onset
$\lambda_i(t)$	Force of infection for susceptibles in group i at time t
β_{ij}	Transmission relative rate from age group j to i
κ	Infectivity ratio of H_i to I_i individuals, ($i = 1, 2, \dots, m$)

In 2007, Feng et al. proposed a general distribution homogeneous model by using the method of probability [24]. Referring to the research method, to investigate the effects of non-exponential stage distributions assumed in heterogeneous epidemic models, we establish an SEIHR model with m subgroups and arbitrarily distributed disease stages.

Hypothesis:

- 1) The total population is divided into m groups, each of which includes S_i, E_i, I_i, H_i, R_i , $N_i = S_i + E_i + I_i + H_i + R_i$, ($i = 1, 2, \dots, m$);
- 2) The initial conditions: $S_i(0) > 0, E_i(0) > 0, I_i(0) > 0, H_i(0) = R_i(0) = 0$.

See other more detailed definitions of variables and parameters listed in Table 1.

The SEIHR model with m subpopulations and arbitrarily distributed disease stages consists of the following integral and differential equations:

$$\begin{aligned}
 \frac{dS_i}{dt} &= -\lambda_i(t)S_i(t), \\
 E_i &= \int_0^t \lambda_i(s)S_i(s)Q_i(t-s)ds + E_i(0)Q_i(t), \\
 I_i &= \int_0^t \int_0^\tau \lambda_i(s)S_i(s)g_{Q_i}(\tau-s)dsP_i(t-\tau)L_i(t-\tau)d\tau + I_i(0)P_i(t)L_i(t) \\
 &\quad + \int_0^t E_i(0)g_{Q_i}(s)P_i(t-s)L_i(t-s)ds, \\
 H_i &= \int_0^t \int_0^\tau \lambda_i(s)S_i(s)g_{Q_i}(\tau-s)dsP_i(t-\tau)[1 - L_i(t-\tau)]d\tau + I_i(0)P_i(t)[1 - L_i(t)] \\
 &\quad + \int_0^t E_i(0)g_{Q_i}(s)P_i(t-s)[1 - L_i(t-s)]ds, \\
 R_i &= \int_0^t \int_0^\tau \int_0^\delta \lambda_i(s)S_i(s)g_{Q_i}(\delta-s)ds g_{P_i}(\tau-\delta)d\delta d\tau + \int_0^t I_i(0)g_{P_i}(\tau)d\tau \\
 &\quad + \int_0^t \int_0^\tau E_i(0)g_{Q_i}(s)g_{P_i}(\tau-s)ds d\tau \quad (i = 1, 2, \dots, m),
 \end{aligned} \tag{2.1}$$

where

$$g_{Q_i}(s) = -\dot{Q}_i(s), \quad g_{P_i}(s) = -\dot{P}_i(s),$$

and

$$\lambda_i(t) = \sum_{j=1}^m \beta_{ij}(I_j(t) + \kappa H_j(t)).$$

The latency in the system (2.1) considered above is the general distribution. Next, we consider exponential distribution and gamma distribution instead of general distribution to derive two ODE models with different distributions of the latency period, respectively.

2.2. Model with exponential distribution of latency (EDM)

Let $Q_i(t) = e^{-\alpha t}, P_i(t) = e^{-\mu t}, L_i(t) = e^{-\gamma t}$, then $g_{Q_i}(t) = \alpha e^{-\alpha t}, g_{P_i}(t) = \mu e^{-\mu t}, g_{L_i}(t) = \gamma e^{-\gamma t}$.

Referring to Feng et al. [18, 25, 30], system (2.1) can be simplified to the following ODE model (see Figure 1 for flowchart):

$$\begin{aligned}\frac{dS_i}{dt} &= -\lambda_i(t)S_i(t), \\ \frac{dE_i}{dt} &= \lambda_i(t)S_i(t) - \alpha E_i(t), \\ \frac{dI_i}{dt} &= \alpha E_i(t) - (\mu + \gamma)I_i(t), \\ \frac{dH_i}{dt} &= \gamma I_i(t) - \mu H_i(t), \\ \frac{dR_i}{dt} &= \mu I_i(t) + \mu H_i(t) \quad (i = 1, 2, \dots, m).\end{aligned}\tag{2.2}$$

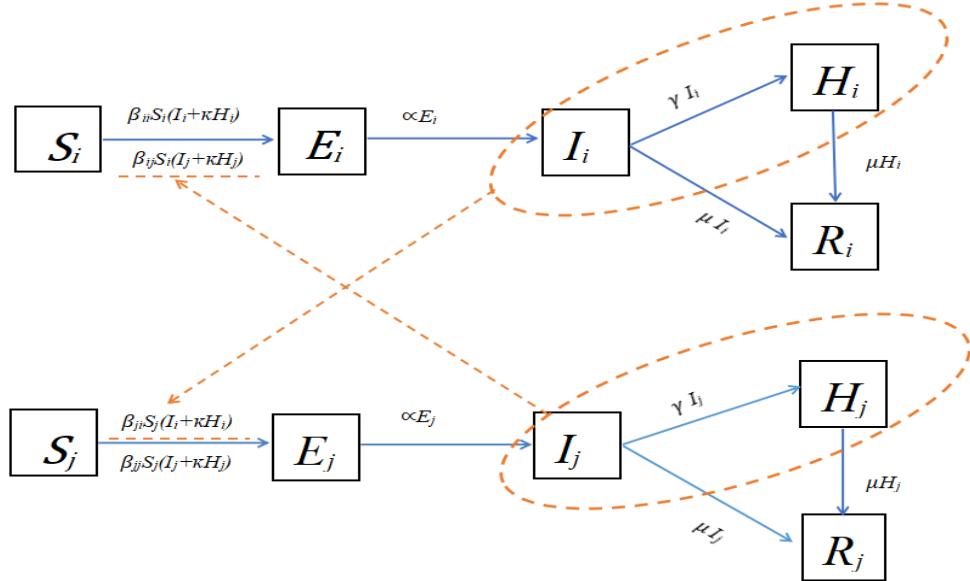


Figure 1. A transition diagram of model (2.2), which is degenerated from model (2.1) when the latency period follows an exponential distribution (i, j denote different subpopulations).

Referring to [4, 29], the basic reproduction number is obtained by the next generation matrix method. The increasing rate of secondary infection and disease progress in disease compartment i are denoted by $\mathcal{F}_i, \mathcal{V}_i$, ($i = 1, 2, \dots, m$), respectively.

$$\mathcal{F}_i = \begin{bmatrix} S_i(t) \sum_{j=1}^m \beta_{ij} (I_j(t) + \kappa H_j(t)) \\ 0 \\ 0 \end{bmatrix}, \quad \mathcal{V}_i = \begin{bmatrix} \alpha E_i(t) \\ (\mu + \gamma) I_i(t) - \alpha E_i(t) \\ \mu H_i(t) - \gamma I_i(t) \end{bmatrix}.$$

The disease-free equilibrium of system (2.2) is

$$E_a = (N_1, 0, 0, \dots, N_i, 0, 0, \dots, N_m, 0, 0, 0, 0).$$

The Jacobi matrix of $\mathcal{F}_i, \mathcal{V}_i$ at E_a is

$$F_i = \begin{bmatrix} 0 & \beta_{i1}N_i & \kappa\beta_{i1}N_i & \cdots & 0 & \beta_{im}N_i & \kappa\beta_{im}N_i \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix}_{3 \times 3m},$$

$$V_i = \begin{bmatrix} 0 & 0 & 0 & \cdots & \alpha & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & -\alpha & \mu + \gamma & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & -\gamma & \mu & \cdots & 0 & 0 & 0 \end{bmatrix}_{3 \times 3m}.$$

\mathcal{F}, \mathcal{V} are defined as the increasing rate of secondary infection and disease progress in the total population, respectively:

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_1 \\ \vdots \\ \mathcal{F}_m \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \mathcal{V}_1 \\ \vdots \\ \mathcal{V}_m \end{bmatrix},$$

and

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix}, \quad V = \begin{bmatrix} V_1 \\ \vdots \\ V_m \end{bmatrix}.$$

The next generation regeneration matrix is FV^{-1} , the basic reproduction number is defined as follows:

$$R_0 = \rho(FV^{-1}),$$

where ρ denotes the spectral radius of a matrix.

2.3. Model with gamma distribution of latency(GDM)

Let $Q_i(t) = \sum_{q=1}^n \frac{(nat)^{q-1} e^{-nat}}{(q-1)!}$, $P_i(t) = e^{-\mu t}$, $L_i(t) = e^{-\gamma t}$, then $g_{Q_i}(t) = \frac{n\alpha(nat)^{n-1}}{(n-1)!} e^{-nat}$, $g_{L_i}(t) = \gamma e^{-\gamma t}$.

Referring to Feng et al. [18, 25, 30], system (2.1) can be simplified to the following ODE model (see Figure 2 for flowchart):

$$\begin{aligned} \frac{dS_i}{dt} &= -\lambda_i(t)S_i(t), \\ \frac{dE_i^1}{dt} &= \lambda_i(t)S_i(t) - n\alpha E_i^1(t), \\ \frac{dE_i^q}{dt} &= n\alpha E_i^{q-1}(t) - n\alpha E_i^q(t), \quad q = 2, 3, \dots, n, \\ \frac{dI_i}{dt} &= n\alpha E_i^n(t) - (\mu + \gamma)I_i(t), \\ \frac{dH_i}{dt} &= \gamma I_i(t) - \mu H_i(t), \\ \frac{dR_i}{dt} &= \mu I_i(t) + \mu H_i(t) \quad (i = 1, 2, \dots, m). \end{aligned} \tag{2.3}$$

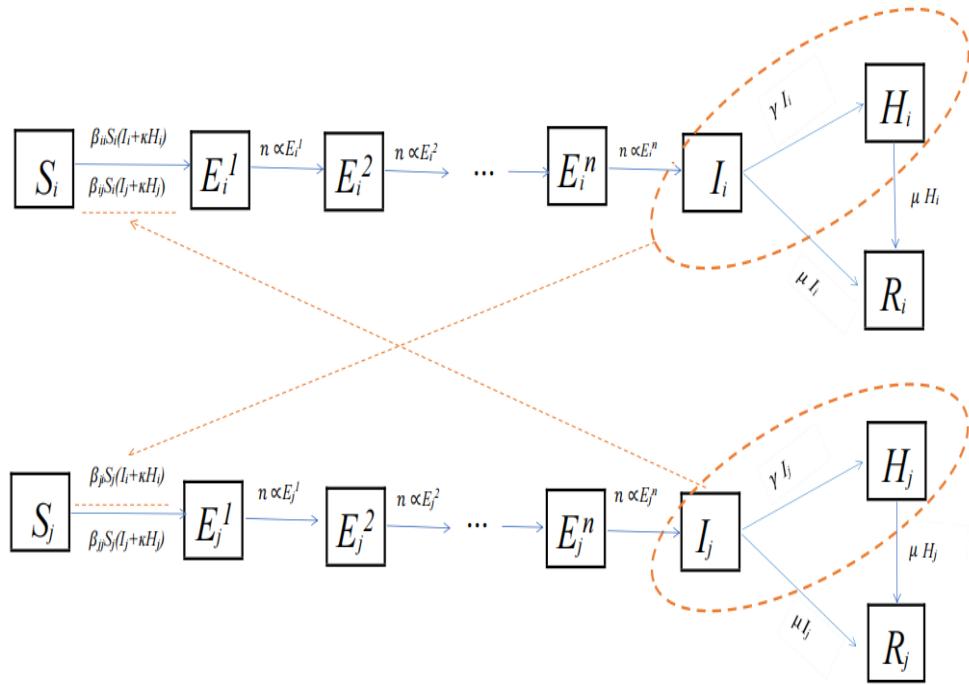


Figure 2. A transition diagram of model (2.3), which is degenerated from model (2.1) when the latency period follows a gamma distribution (i, j denote different subpopulations).

Referring to [4, 29], the basic reproduction number is obtained by the next generation matrix method. the increasing rate of secondary infection and disease progress in disease compartment i are denoted by $\mathcal{F}_i, \mathcal{V}_i$, ($i = 1, 2, \dots, m$), respectively.

$$\mathcal{F}_i = \begin{bmatrix} S_i(t) \sum_{j=1}^m \beta_{ij} (I_j(t) + \kappa H_j(t)) \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{(n+2)}, \quad \mathcal{V}_i = \begin{bmatrix} n\alpha E_i^1(t) \\ n\alpha E_i^2(t) - n\alpha E_i^1(t) \\ \vdots \\ n\alpha E_i^n(t) - n\alpha E_i^{n-1}(t) \\ (\mu + \gamma)I_i(t) - \alpha E_i^n(t) \\ \mu H_i(t) - \gamma I_i(t) \end{bmatrix}_{(n+2)}.$$

The disease-free equilibrium of system (2.3) is

$$E_b = (N_1, 0, 0, \dots, N_i, 0, 0, \dots, N_m, 0, 0, \dots, 0).$$

The Jacobi matrix of $\mathcal{F}_i, \mathcal{V}_i$ at E_b is

$$F_i = \begin{bmatrix} 0 & \dots & 0 & \beta_{i1} N_i & \kappa \beta_{i1} N_i & \dots & 0 & \dots & 0 & \beta_{im} N_i & \kappa \beta_{im} N_i \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & \dots & 0 & 0 & 0 \end{bmatrix}_{(n+2) \times (m(n+2))},$$

$$V_i = \begin{bmatrix} 0 & \cdots & 0 & n\alpha & 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & -n\alpha & n\alpha & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & -n\alpha & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & -n\alpha & n\alpha & 0 & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & -n\alpha & \mu + \gamma & 0 & 0 & \cdots & 0 \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & -\gamma & \mu & 0 & \cdots & 0 \end{bmatrix}_{(n+2) \times (m(n+2))}.$$

\mathcal{F}, \mathcal{V} are defined as the increasing rate of secondary infection and disease progress in the total population, respectively:

$$\mathcal{F} = \begin{bmatrix} \mathcal{F}_1 \\ \vdots \\ \mathcal{F}_m \end{bmatrix}, \quad \mathcal{V} = \begin{bmatrix} \mathcal{V}_1 \\ \vdots \\ \mathcal{V}_m \end{bmatrix},$$

and

$$F = \begin{bmatrix} F_1 \\ \vdots \\ F_m \end{bmatrix}, \quad V = \begin{bmatrix} V_1 \\ \vdots \\ V_m \end{bmatrix},$$

The next generation regeneration matrix is FV^{-1} , the basic reproduction number is defined as follows:

$$R_0 = \rho(FV^{-1}),$$

where ρ denotes the spectral radius of a matrix.

3. Comparison of EDM and GDM based on real data

In this section, we divide the population of Guangdong Province into four age groups based on the transmission characteristics of COVID-19: group 1, those ≤ 5 years old; group 2, those 6–19 years old; group 3, those 20–64 years old; and group 4, those ≥ 65 years old, and conduct global uncertainty and sensitivity analysis through Latin hypercube sampling (LHS) and Partial Rank Correlation Coefficients (PRCC). Then, according to Table 2, we used numerical simulations to explore the discrepancies between EDM and GDM in revealing disease transmission and assessing control strategies.

3.1. Uncertainty and sensitivity analysis

Sensitivity analysis provides significant information concerning how uncertainty and variability of model parameters may impact model consequences and which parameters are most influential. In this subsection, we perform uncertainty analysis of the input data for all parameters of the model mainly by the LHS method and further use PRCC to investigate the global sensitivity of the related parameters [31–34]. In this way, it is determined which factors are important in the disease epidemic for better selection of epidemic control measures.

Figure 3 depicts the correlation between parameters and the total number of infected persons over time, and the gray area in the figure represents PRCCs that are not significantly distinct from zero. Note that the PRCC values range from -1 to 1 . Negative (positive) values represent a negative (positive)

correlation of the parameter with the model output. A negative (positive) correlation suggests that a negative (positive) variation of the parameter will decrease (increase) the output of the model. From the Figure 3, we clearly notice that, for both EDM and GDM, the correlation between parameters and total infections showed consistent results: the most significant positive effect on the number of total infected persons is β_{33} , κ . That is, when the transmission rate β_{33} between the third age group and the hospitalization infection rate κ increase, it will make the total number of infected individuals increase rapidly, and more people will be infected. Next, the correlation between transmission rate and total number of infected persons varies by age group, and it can be noted that an increase in the transmission rate of the third age group has a greater impact on the increase in the total number of infections than the other three age groups.

In addition, the correlation of parameters may change over time. Especially, the hospitalization rate γ have a more positive correlation with the output of the target model, which is the total number of infected persons at the beginning of the disease. With the progression of the disease and the increase in the total number of infected individuals, this parameter will negatively affect the target output. In fact, due to the initial lack of understanding of the characteristics and mode of infection of the disease, the flow of infected people is promoted as hospitalization rates continued to increase, increasing the number of people who contracted the disease. However, as time passes, much progress has been made in the understanding of how diseases transmitted and cured, resulting in fewer people being infected. And α , μ show a significant negative impact on the total number of infected persons over time. Higher progression rate of exposed individuals to infectives (α) and recovery rates (μ) directly imply a shorter latency and recovery period, which further suggests that the time to make effective contacts are less, thereby reduction in the spread of infection and thus the number of infections.

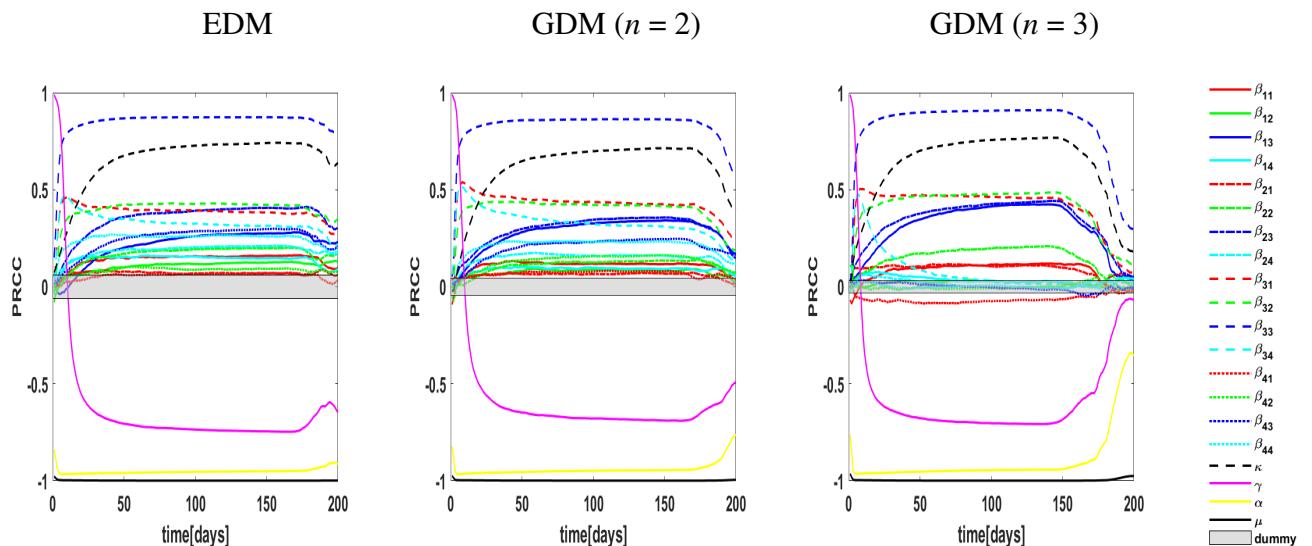


Figure 3. Time-varying PRCCs sensitivity indexes of total infectives in three differential models.

Therefore, based on the results of sensitivity analysis, it is indicated that decreasing the transmission rate and the hospitalization infection rate or increasing the hospitalization rate could effectively reduce the number of infected individuals. Accordingly, we can take the following control measures: reducing contact between populations by isolating at home and wearing masks when traveling and

the hospitalization infection rate by taking preventive measures such as sterilization and isolation in hospitals, increasing the hospitalization rate by shortening the period of diagnosis, and improving the medical system.

3.2. Effectiveness of different interventions on EDM and GDM

In this subsection, we perform numerical simulations for three models, including EDM and GDM ($n = 2, 3$), to compare the results when the models are used to estimate different interventions specified by γ (hospitalization rate) and/or κ (hospitalization infection rate), and then investigate how the models may obtain different assessments on the effectiveness of various control measures. Different control strategies are presented separately for Guangdong province, and some key indicators are used to evaluate the effectiveness of the control strategies include the peak value, peak time, final size of an outbreak and the infection eradication time* [36]. The simulation results are shown in Figures 4–6.

Figure 4 shows the numerical simulation of EDM, GDM ($n = 2$) and GDM ($n = 3$) based on the parameters of Guangdong province, it plots active infectious individuals (red curve), cumulative infections (blue curve). The four rows compare for scenarios based on implementing interventions such as performing hospital sterilization to reduce the hospitalization infection rate (κ) and shortening the diagnosis period to increase the hospitalization rate (γ). It consists of a baseline scenario and three strategies. We observe from Figure 4 that when hospitalization rate (γ) increases (baseline scenario and strategy I) or the hospitalization infection rate (κ) decreases (baseline scenario and Strategy II), the final size and peak size will be reduced and the peak time will be advanced in all three models (see rows 1–3). By comparing strategy III with Strategies I and II (see rows 2–4), it can be observed that the most effective way to control infectious diseases is to combine the two methods (Strategy III). See Table 3 for different strategies.

A more detailed comparison of the infection eradication time, the peak value and time for the three control strategies is shown in Figure 5. In particular, comparing Strategies I, II and III to the baseline scenario, they all lead to earlier end of infection, lower peak value and earlier peak time. By comparing Figure 5(a) with Figure 5(b),(c), we found that the decrease in hospitalization infection rate has a more pronounced effect on the end time of infection than on the peak value and time, but increasing hospitalization rate has a similar effect on time to infection eradication, peak value and time. What's more, we observe that all models provided consistent assessments of control strategies, but we also note that EDM has a later infection eradication time, smaller peak value and earlier peak time, while GDM, specifically GDM ($n = 3$) has an earlier infection eradication time, larger peak value and later peak time.

It is clear from Figure 5(a) and Figure 6 that EDM, GDM ($n = 2$) and GDM ($n = 3$) appear to have the same final size in the case of the same control strategy, but the infection eradication time (i.e., the time to reach the final size) can be significantly different. By comparing three strategies with the baseline scenario, decrease in hospitalization infection rate or increase in hospitalization rate is found to be effective in reducing the final disease size as well as shortening the end time of the outbreak.

*The infection eradication time is denoted as T , $T = \inf \{t \in R_+ \mid I(t) = \varepsilon, 0 < \varepsilon < 1\}$, choose the initial number of infected units $I(0)$ strictly greater than ε , T is the first time at which the state variable I drops to ε . From a biological point of view, by setting $\varepsilon < 1$ as the condition to identify the eradication time means assuming the infection goes extinct when there is less than one infectious individual in the population [35].

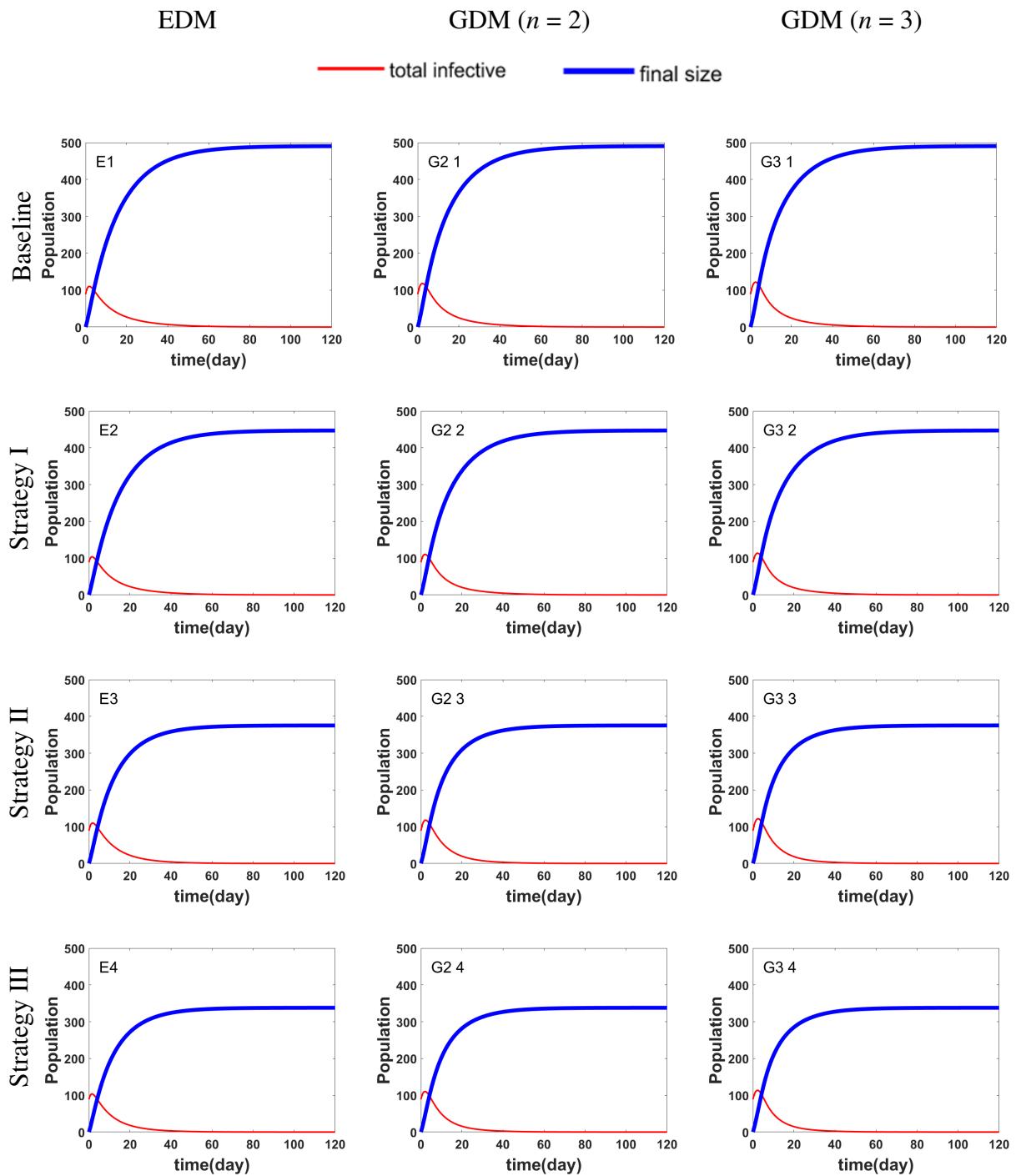


Figure 4. Comparison of the epidemic sizes generated by the models EDM, GDM ($n = 2$) and GDM ($n = 3$) under the baseline scenario (top row) and strategies I, II and III (rows 2, 3 and 4, respectively).

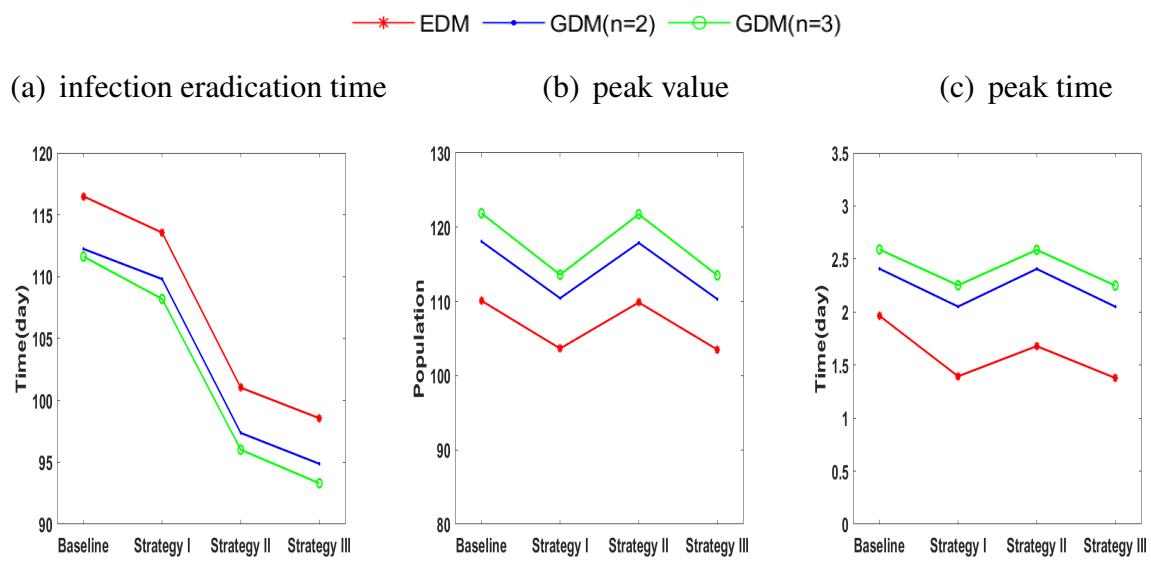


Figure 5. Comparison of infection eradication time (a), peak value (b) and peak time (c) generated by the three models under the baseline scenario and the three control strategies I, II and III.

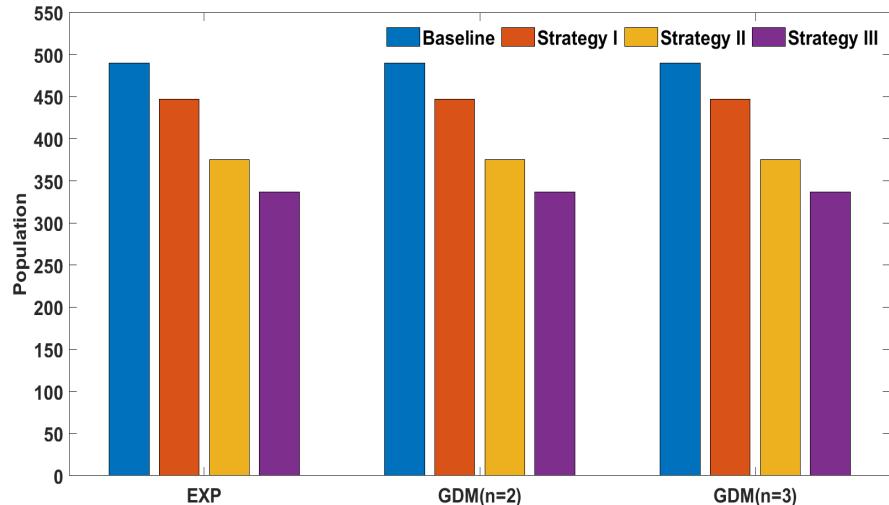


Figure 6. Comparison of the final size generated by the three models under the baseline scenario and the three control strategies I, II and III, in Guangdong Province.

3.3. A more detailed comparison of EDM and GDM on interventions

To further investigate the difference between EDM and GDM with the change of control parameters, we simulate contour plots of the infection eradication time, peak value, peak time with respect to hospitalization infection rate (κ) and hospitalization rate (γ) in three different models, including EDM and GDM ($n = 2, 3$).

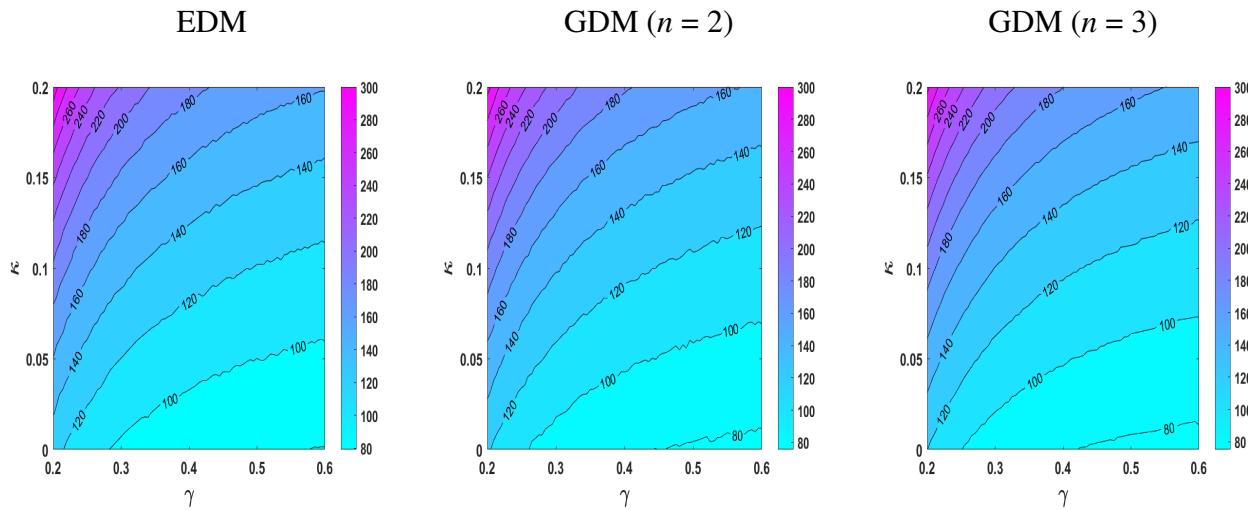


Figure 7. Contour plots of the infection eradication time with respect to κ (hospitalization infection rate) and γ (hospitalization rate) in three different models.

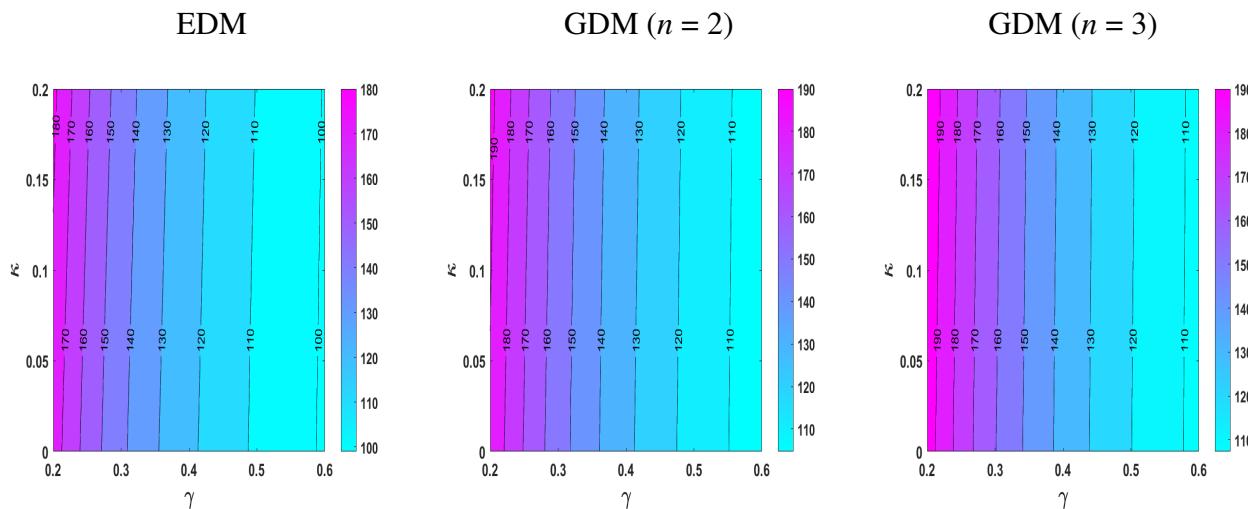


Figure 8. Contour plots of the peak value of the total number of infected individuals with respect to κ (hospitalization infection rate) and γ (hospitalization rate) in three different models.

Figures 7–9 further confirm the conclusions in Figures 4 and 5: in both EDM and GDM, as γ increases and κ decreases, the peak time and infection elimination time are advanced, and the peak size is lowered. This suggests that control measures such as disinfection and isolation in hospitals and a shorter diagnosis period during the stage of the disease outbreak would reduce the severity of the epidemic. Figures 7 and 8 indicate that EDM overestimates the infection end time and underestimates the infection peak value when compared to GDM. Figure 9 shows that EDM and GDM are inconsistent in their estimates of the peak time: (a) if $\gamma < 0.24$, EDM overestimates peak time compared to GDM; (b) if $0.24 \leq \gamma \leq 0.2625$, when the hospitalization infection rate is higher and the hospitalization rate is lower (i.e., Top left corner in Figure 9 (b)), EDM overestimates the peak time compared to GDM; when the hospitalization infection rate is low and the hospitalization rate is higher (i.e., Bottom right

corner in Figure 9(b)), EDM underestimates the peak time compared to GDM; (c) if $\gamma > 0.265$, EDM underestimates the peak time compared to GDM.

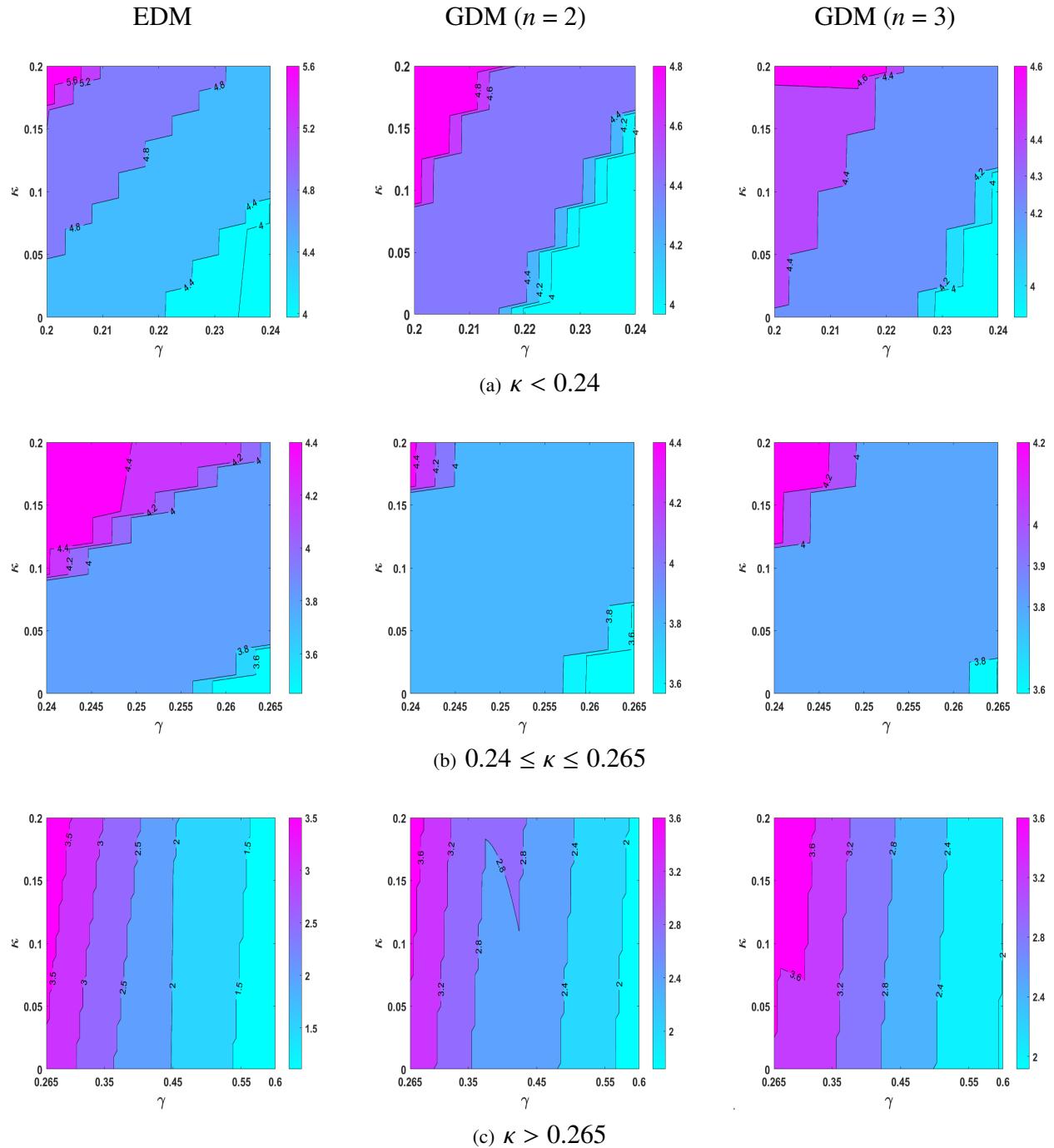


Figure 9. Contour plots of the peak time of the total number of infected individuals with respect to κ (hospitalization infection rate) and γ (hospitalization rate) in three different models.

Table 2. Parameter values in Guangdong Province.

Parameters	Values	Unit	Definitions	Reference
β_{11}	4.81×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate among age group 1	[9]
β_{12}	6.20×10^{-10}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 2 to 1	[9]
β_{13}	2.20×10^{-10}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 3 to 1	[9]
β_{14}	1.01×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 4 to 1	[9]
β_{21}	2.31×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 1 to 2	[9]
β_{22}	2.43×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate among age group 2	[9]
β_{23}	3.30×10^{-10}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 3 to 2	[9]
β_{24}	1.82×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 4 to 2	[9]
β_{31}	1.53×10^{-8}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 1 to 3	[9]
β_{32}	4.83×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 2 to 3	[9]
β_{33}	1.25×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate among age group 3	[9]
β_{34}	3.42×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 4 to 3	[9]
β_{41}	4.25×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 1 to 4	[9]
β_{42}	2.45×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 2 to 4	[9]
β_{43}	4.00×10^{-10}	$1/(\text{individual} \times \text{day})$	Transmission relative rate from age group 3 to 4	[9]
β_{44}	4.83×10^{-9}	$1/(\text{individual} \times \text{day})$	Transmission relative rate among age group 4	[9]
$\frac{1}{\alpha}$	5.2	day	Latent period	[9]
$\frac{1}{\gamma}$	2.0383	day	Mean time from onset to hospitalization	[9]
κ	0.091	day	Infectivity ratio of H_i to I_i individuals, ($i = 1, 2, \dots, m$)	[37]
$S_1(0)$	7,153,800	persons	Initial number of S_1	[9]
$S_2(0)$	7,153,800	persons	Initial number of S_2	[9]
$S_3(0)$	78,760,477	persons	Initial number of S_3	[9]
$S_4(0)$	10,664,042	persons	Initial number of S_4	[9]
$E_1(0)$	16	persons	Initial number of E_1	[9]
$E_2(0)$	31	persons	Initial number of E_2	[9]
$E_3(0)$	317	persons	Initial number of E_3	[9]
$E_4(0)$	64	persons	Initial number of E_4	[9]
$I_1(0)$	2	persons	Initial number of I_1	[9]
$I_2(0)$	2	persons	Initial number of I_2	[9]
$I_3(0)$	68	persons	Initial number of I_3	[9]
$I_4(0)$	17	persons	Initial number of I_4	[9]
$H_1(0)$	2	persons	Initial number of H_1	[9]
$H_2(0)$	3	persons	Initial number of H_2	[9]
$H_3(0)$	109	persons	Initial number of H_3	[9]
$H_4(0)$	24	persons	Initial number of H_4	[9]

Table 3. Three control strategies in Guangdong Province.

Baseline scenario	$\kappa = 0.091$	$\gamma = 1/2.0383$
Strategy I	$\kappa = 0.091$	$\gamma = 0.55$
Strategy II	$\kappa = 0.05$	$\gamma = 1/2.0383$
Strategy III	$\kappa = 0.05$	$\gamma = 0.55$

4. Conclusions and discussion

In this paper, we have developed and investigated two SEIHR epidemic models with exponential and gamma distribution of latency, taking into account the effect of population heterogeneity and infectious hospitalized individuals. Firstly, we have theoretically derived the formula for the basic reproduction number R_0 . Secondly, we have performed sensitivity analysis of the model parameters using the partial rank correlation coefficient (PRCC) approach and identified key parameters affecting disease transmission to derive reasonable control measures. Next, we have evaluated the impact of EDM and GDM on interventions based on age heterogeneity. Finally, using the parameters of COVID-19 in Guangdong Province, we have obtained the similarities and differences between EDM and GDM in revealing disease transmission and assessing control strategies through numerical simulation results.

Our results suggest that the evaluation of the different strategies by EDM and GDM is synchronous, however, when control strategies are the same, it seems that GDM may not have an effect on the overall size of an epidemic but it does affect the process of infection and further change the overall duration of an outbreak. EDM may result in later infection eradication time and lower peak value compared to GDM. Moreover, another finding of our study is that EDM and GDM produce inconsistent results for predicting peak time, and the outcome of the prediction may depend on the strength of the control strategy.

When we consider a mathematical model with population heterogeneity to describe disease transmission, we underestimate the infection eradication time and overestimate the peak value, assuming that the latent period follows a gamma distribution (actually an exponential distribution). In contrast, when the latent period is assumed to follow an exponential distribution (actually a gamma distribution), we overestimate the infection eradication time and underestimate the peak value. Furthermore, the strength of the control strategy (especially considering the effect of hospitalization rate) may also have an impact on the prediction results of the peak time when the exponential distribution of latency is replaced by a gamma distribution. Specifically, if the control is weak, the peak time may be underestimated, and if the control is strong, the peak time may be overestimated. Therefore, when applying mathematical models to describe the spread of diseases, one must be careful to make assumptions about the distribution of latency periods. If unreasonable assumptions are made about the distribution of latency periods, we may overestimate or underestimate the peak value, peak time, and infection eradication time. Meanwhile, estimating peak time needs to take into account the effect of control strategies when considering different distributions of latency periods.

In contrast to the existing literatures, it should be noted that the main innovation of this paper is perhaps the first consideration of the difference between latency obeying exponential and non-exponential distributions in a heterogeneous model. In the literature [27], the model comparisons done in Blyuss et al. focus on the differences between exponential and non-exponential distributions of homogeneous models. In this paper, our results show that the estimates for the end time of infection are consistent with the results obtained by Blyuss et al. (i.e., increasing the number of stages in the latency period leads to an earlier overall end time), but not for the time to peak, which may depend on the intensity of control measures. Moreover, our results suggest that in the heterogeneous model, increasing the number of stages of latency leads to an increase in the number of infections. These results significantly expand the understanding of the law of heterogeneous infectious disease. The model proposed in this paper is obtained under the assumption of ignoring demographic dynamics. However, for example,

AIDS, tuberculosis, and other such long-term epidemics, we should not ignore the influence of demographic dynamics and therefore need to consider long time scales models of heterogeneous infectious diseases. Moreover, in our paper, the results obtained by comparing the exponential and gamma distributions help to understand the disease transmission pattern more accurately, but for the latent period, there are not only exponential and gamma distributions [38]. Hence, it is necessary to consider models of other distributions and analyze their differences in understanding disease transmission. We leave it to future work to address these questions.

Acknowledgments

The authors express their sincere gratitude to Yan Chen, Qinjin Luo, Nannan Li, and Lirong Guo for their valuable input in the numerical simulation section of this paper. Their insights greatly enhanced the quality of our work.

Furthermore, the authors would like to extend their appreciation to the three anonymous reviewers whose meticulous review and astute suggestions played an instrumental role in enhancing the quality of this manuscript, leading to a significant improvement.

This research is supported by the National Natural Science Foundation of China (No. 12271401, 11871179) and the Natural Science Foundation of Tianjin, China (No. 22JCYBJC00080).

Conflict of interest

The authors declare there is no conflict of interest.

References

1. P. M. Martin, E. Martin-Granel, 2500-year evolution of the term epidemic, *Emerging Infect. Dis.*, **12** (2006), 976–980. <https://doi.org/10.3201/eid1206.051263>
2. F. Brauer, C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology*, Springer, 2012.
3. X. Wang, S. Wang, J. Wang, L. Rong, A multiscale model of COVID-19 dynamics, *Bull. Math. Biol.*, **84** (2022), 1–41. <https://doi.org/10.1007/s11538-022-01058-8>
4. J. Cui, Y. Wu, S. Guo, Effect of non-homogeneous mixing and asymptomatic individuals on final epidemic size and basic reproduction number in a meta-population model, *Bull. Math. Biol.*, **84** (2022), 1–22. <https://doi.org/10.1007/s11538-022-00996-7>
5. Z. Ma, Y. Zhou, J. Wu, *Mathematical Modeling and Research on the Dynamics of Infectious Diseases*, Science Press, 2004.
6. J. Cui, Y. Zhang, Z. Feng, Influence of non-homogeneous mixing on final epidemic size in a meta-population model, *J. Biol. Dyn.*, **13** (2019), 31–46. <https://doi.org/10.1080/17513758.2018.1484186>
7. A. Lajmanovich, J. A. Yorke, A deterministic model for gonorrhea in a nonhomogeneous population, *Math. Biosci.*, **28** (1976), 221–236. [https://doi.org/10.1016/0025-5564\(76\)90125-5](https://doi.org/10.1016/0025-5564(76)90125-5)

8. A. Nold, Heterogeneity in disease-transmission modeling, *Math. Biosci.*, **52** (1980), 227–240. [https://doi.org/10.1016/0025-5564\(80\)90069-3](https://doi.org/10.1016/0025-5564(80)90069-3)
9. G. Wang, S. Zhang, J. Yang, X. Xu, Study of coupling the age-structured contact patterns to the COVID-19 pandemic transmission, *Acta Phys. Sin.*, **70** (2021), 210–220. <https://doi.org/10.7498/aps.70.20201371>
10. G. Dimarco, B. Perthame, G. Toscani, M. Zanella, Kinetic models for epidemic dynamics with social heterogeneity, *J. Math. Biol.*, **81** (2021), 4. <https://doi.org/10.1007/s00285-021-01630-1>
11. J. F. David, Epidemic models with heterogeneous mixing and indirect transmission, *J. Biol. Dyn.*, **12** (2018), 375–399. <https://doi.org/10.1080/17513758.2018.1467506>
12. Z. Feng, A. N. Hill, A. T. Curns, J. W. Glasser, Evaluating targeted interventions via meta-population models with multi-level mixing, *Math. Biosci.*, **287** (2017), 93–104. <https://doi.org/10.1016/j.mbs.2016.09.013>
13. J. W. Glasser, Z. Feng, S. B. Omer, P. J. Smith, L. E. Rodewald, The effect of heterogeneity in uptake of the measles, mumps, and rubella vaccine on the potential for outbreaks of measles: a modelling study, *Lancet Infect. Dis.*, **16** (2016), 599–605. [https://doi.org/10.1016/S1473-3099\(16\)0004-9](https://doi.org/10.1016/S1473-3099(16)0004-9)
14. D. J. Rodríguez, L. Torres-Sorando, Models of infectious diseases in spatially heterogeneous environments, *Bull. Math. Biol.*, **63** (2001), 547–571. <https://doi.org/10.1006/bulm.2001.0231>
15. L. Sattenspiel, K. Dietz, A structured epidemic model incorporating geographic mobility among regions, *Math. Biosci.*, **128** (1995), 71–91. [https://doi.org/10.1016/0025-5564\(94\)00068-B](https://doi.org/10.1016/0025-5564(94)00068-B)
16. M. Zanella, C. Bardelli, G. Dimarco, S. Deandrea, P. Perotti, M. Azzi, et al., A data-driven epidemic model with social structure for understanding the COVID-19 infection on a heavily affected Italian Province, *Math. Models Methods Appl. Sci.*, **31** (2021), 2533–2570. <https://doi.org/10.1142/S021820252150055X>
17. H. J. Wearing, P. Rohani, M. J. Keeling, Appropriate models for the management of infectious diseases, *PLoS Med.*, **2** (2005), e174. <https://doi.org/10.1371/journal.pmed.0020174>
18. Z. Feng, X. Wang, Y. Shi, J. Cui, Evaluations of interventions using mathematical models with exponential and non-exponential distributions for disease stages: the case of Ebola, *Bull. Math. Biol.*, **79** (2017), 2149–2173. <https://doi.org/10.1007/s11538-017-0324-z>
19. Y. Deng, C. You, Y. Liu, J. Qin, X. Zhou, Estimation of incubation period and generation time based on observed length-biased epidemic cohort with censoring for COVID-19 outbreak in China, *Biometrics*, **77** (2021), 929–941. <https://doi.org/10.1111/biom.13325>
20. S. A. Lauer, K. H. Grantz, Q. Bi, F. K. Jones, Q. Zheng, H. R. Meredith, et al., The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application, *Ann. Intern. Med.*, **172** (2020), 577–582. <https://doi.org/10.7326/M20-0504>
21. X. Pan, Y. Chen, A. Wang, J. Wang, L. Ye, S. Gu, et al., Study on transmission dynamic of 15 clusters of COVID-2019 cases in Ningbo, *Chin. J. Epidemiol.*, **41** (2020), 2010–2014. <http://dx.doi.org/10.3760/cma.j.cn112338-20200330-00466>

22. L. C. Tindale, J. E. Stockdale, M. Coombe, E. S. Garlock, W. Y. V. Lau, M. Saraswat, et al., Evidence for transmission of COVID-19 prior to symptom onset, *eLife*, **9** (2020), e57149. <https://doi.org/10.7554/eLife.57149>

23. Z. Du, J. Gu, J. Li, X. Lin, Y. Wang, L. Chen, et al., Estimating the distribution of COVID-19 incubation period by interval-censored data estimation method, *Chin. J. Epidemiol.*, **41** (2020), 1000–1003. <http://dx.doi.org/10.3760/cma.j.cn112338-20200313-00331>

24. Z. Feng, D. Xu, H. Zhao, Epidemiological models with non-exponentially distributed disease stages and applications to disease control, *Bull. Math. Biol.*, **69** (2007), 1511–1536. <https://doi.org/10.1007/s11538-006-9174-9>

25. Z. Feng, Y. Zheng, N. Hernandez-Ceron, H. Zhao, J. W. Glasser, A. N. Hill, Mathematical models of Ebola-Consequences of underlying assumptions, *Math. Biosci.*, **277** (2016), 89–107. <https://doi.org/10.1016/j.mbs.2016.04.002>

26. M. A. Capistran, A. Capella, J. A. Christen, Forecasting hospital demand in metropolitan areas during the current COVID-19 pandemic and estimates of lockdown-induced 2nd waves, *PLoS One*, **16** (2021), e0245669. <https://doi.org/10.1371/journal.pone.0245669>

27. K. B. Blyuss, Y. N. Kyrychko, Effects of latency and age structure on the dynamics and containment of COVID-19, *J. Theor. Biol.*, **513** (2021), 110587. <https://doi.org/10.1016/j.jtbi.2021.110587>

28. Y. Chen, H. Song, S. Liu, Evaluations of COVID-19 epidemic models with multiple susceptible compartments using exponential and non-exponential distribution for disease, *Inf. Dis. Modell.*, **7** (2022), 795–810. <https://doi.org/10.1016/j.idm.2022.11.004>

29. P. Van den Driessche, J. Watmough, Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission, *Math. Biosci.*, **180** (2002), 29–48. [https://doi.org/10.1016/S0025-5564\(02\)00108-6](https://doi.org/10.1016/S0025-5564(02)00108-6)

30. L. Bolzoni, R. Della Marca, M. Groppi, On the optimal control of SIR model with Erlang-distributed infectious period: isolation strategies, *J. Math. Biol.*, **83** (2021), 1–21. <https://doi.org/10.1007/s00285-021-01668-1>

31. S. Marino, I. B. Hogue, C. J. Ray, D. E. Kirschner, A methodology for performing global uncertainty and sensitivity analysis in systems biology, *J. Theor. Biol.*, **254** (2008), 178–196. <https://doi.org/10.1016/j.jtbi.2008.04.011>

32. H. Song, L. Guo, Z. Jin, S. Liu, Modelling and stability analysis of ASFV with swill and the virus in the environment, *Math. Biosci. Eng.*, **19** (2022), 13028–13049. <https://doi.org/10.3934/mbe.2022608>

33. J. Franceschi, A. Medaglia, M. Zanella, On the optimal control of kinetic epidemic models with uncertain social features, preprint, arXiv:2210.09201.

34. G. Albi, L. Pareschi, M. Zanella, Control with uncertain data of socially structured compartmental epidemic models, *J. Math. Biol.*, **82** (2021), 1–41. <https://doi.org/10.1007/s00285-021-01630-1>

35. L. Bolzonia, E. Bonacini, R. Della Marcab, M. Groppi, Optimal control of epidemic size and duration with limited resources, *Math. Biosci.*, **315** (2019), 108232. <https://doi.org/10.1016/j.mbs.2019.108232>

- 36. S. J. Schreiber, S. Huang, J. Jiang, H. Wang, Extinction and quasi-stationarity for discrete-time, endemic SIS and SIR models, *SIAM J. Appl. Math.*, **81** (2021), 2195–2217. <https://doi.org/10.1137/20M1339015>
- 37. L. Masandawa, S. S. Mirau, I. S. Mbalawata, J. N. Paul, K. Kreppel, O. M. Msamba, Modeling nosocomial infection of COVID-19 transmission dynamics, *Results Phys.*, **37** (2022), 105503. <https://doi.org/10.1016/j.rinp.2022.105503>
- 38. S. Wan, J. Liu, M. Liu, Progress on the incubation period of COVID-19, *Chin. Sci. Bull.*, **66** (2021), 802–1811. <https://doi.org/10.1360/TB-2020-0698>



AIMS Press

© 2023 the Author(s), licensee AIMS Press. This is an open access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0>)