Journal of Applied Mathematics and Physics, 2026, 14(9), 3551-3570
https://www.scirp.org/journal/jamp
ISSN Online: 2327-4379
ISSN Print: 2327-4352

Optimal Control and Numerical Simulation of a SIHR Model with Nonlinear Birth and Incidence Rates

Qi Zhang

School of Science, Tianjin University of Technology and Education, Tianjin, China
Email: 2501679309@qq.com

How to cite this paper: Zhang, Q. (2026) Optimal Control and Numerical Simulation of a SIHR Model with Nonlinear Birth and Incidence Rates. *Journal of Applied Mathematics and Physics*, **14**, 3551-3570.
https://doi.org/10.4236/jamp.2026.149177

Received: August 20, 2026
Accepted: September 15, 2026
Published: September 18, 2026

Abstract

This paper investigates optimal control strategies for an SIHR epidemic model with a nonlinear birth rate, a nonlinear incidence rate, and limited medical resources. Based on the original model, three control variables are introduced: transmission-suppression control, medical-resource saturation relief control, and treatment-efficiency enhancement control. An optimal control model is then formulated. The existence of an optimal control is established by applying the Fleming and Rishel theorem, and the necessary optimality conditions are derived from Pontryagin's maximum principle. Numerical simulations show that the combined control strategy can substantially reduce the numbers of infected and hospitalized individuals as well as the saturation pressure on medical resources. The results indicate that coordinated interventions and expansion of medical capacity are important for disease prevention and control.

Keywords

Epidemic Model, Optimal Control, Pontryagin's Maximum Principle, Limited Medical Resources, Numerical Simulation

1. Introduction

Epidemic dynamical models are important tools for studying transmission mechanisms, predicting epidemic trends, and evaluating public health interventions. Since Kermack and McKendrick proposed the classical SIR model [1], a large number of epidemic models have been developed and widely used to study disease transmission. To describe more realistically the effects of behavioral changes, increased self-protection awareness, and governmental interventions during an epidemic, Ruan and Wang [2] studied an epidemic model with a nonlinear incidence rate and revealed the influence of nonmonotone incidence on system dynamics. Zhang *et al.*

[3] further incorporated both saturated incidence and a saturated treatment function into an SEIR model and analyzed the stability of equilibria and the influence of limited treatment on disease spread. Chen and Xiang constructed several time-dependent control strategies [4], including transmission control, vaccination, and treatment, for epidemic models with nonlinear incidence rates, and numerical simulations showed that combined control is superior to a single control strategy.

Limited medical resources are also an important factor that cannot be ignored in practical epidemic prevention and control. Zhang and Liu studied an epidemic model with a saturated treatment function [5] and pointed out that limited treatment resources may lead to backward bifurcation. Shan and Zhu further discussed the influence of the number of hospital beds on the complex dynamics of an SIR model [6], showing that medical capacity is an important parameter affecting disease transmission and bifurcation structure. Wang *et al.* [7] considered both individual and public medical resources in a network-based transmission model and analyzed the influence of medical resources on backward bifurcation and optimal control.

By introducing time-dependent control variables, the infection scale and disease burden can be reduced while intervention costs are taken into account, leading to a more economical and effective control scheme. Laarabi *et al.* formulated an optimal control problem for an SIR model with saturated incidence and characterized the optimal control by means of adjoint variables [8]. Lenhart and Workman systematically introduced applications of optimal control theory to biological models and provided a common framework for solving optimality systems by the forward-backward sweep method [9]. Balderrama *et al.* studied optimal control for an SIR model with limited quarantine time [10], and Du *et al.* investigated a stochastic SEIR model with nonlinear incidence and treatment rates, verifying through numerical simulations the inhibitory effect of control variables on disease transmission [11].

However, the optimal control problem for SIHR epidemic models that include both nonlinear incidence and limited medical-resource constraints still requires further investigation. Motivated by this observation, this paper introduces three time-dependent controls into a doubly nonlinear SIHR model: transmission-suppression control, medical-resource saturation relief control, and treatment-efficiency enhancement control. The corresponding optimal control model is then established.

2. Model Formulation

We consider the following epidemic model:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = \left(\dfrac{A}{N}+b\right)N - \dfrac{\beta SI}{\alpha I^2+1} - \mu S \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \dfrac{\beta SI}{\alpha I^2+1} - rI\left(1-\dfrac{H}{k}\right) - v_1 I - \varepsilon_1 I - \mu I \\ \dfrac{\mathrm{d}H}{\mathrm{d}t} = rI\left(1-\dfrac{H}{k}\right) - v_2 H - \varepsilon_2 H - \mu H \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = v_1 I + v_2 H - \mu R \end{cases} . \tag{1}$$

In this study, we consider a general acute infectious disease with hospitalization dynamics. Here, S, I, H and R denote the numbers of susceptible individuals, infectious individuals who have not yet been hospitalized, hospitalized individuals, and recovered individuals at time t, respectively.

The nonlinear birth rate is $\frac{A}{N}+b$, μ is the natural death rate ($b<\mu$), β denotes disease transmissibility, and $\frac{1}{\alpha I^2+1}$ reflects the reduction in transmission caused by protective behavior. The parameter r is the transfer rate from the infected class to hospitalization, k is the maximum hospital admission capacity, v_1 and v_2 are the recovery rates of infected and hospitalized individuals, respectively, and ε_1 and ε_2 are the disease-induced death rates of infected and hospitalized individuals, respectively.

The nonlinear incidence rate considered in the model describes the inhibition of transmission. When the number of infected individuals is small, the incidence rate is close to the βSI; when the number of infected individuals increases, the incidence rate may instead decrease. This reflects the fact that, during a severe epidemic, individuals may spontaneously adopt preventive measures because of heightened risk awareness, social restrictions, or behavioral changes, thereby slowing disease transmission. Such phenomena have been observed in epidemics such as SIRS-type infections and COVID-19.

The basic reproduction number of model (1) at the disease-free equilibrium $E_0=\left(\frac{A}{\mu-b},0,0,0\right)$ is given by $R_0=\frac{\beta A}{(\mu-b)(r+v_1+\varepsilon_1+\mu)}$.

If $R_0<1$, then the disease-free equilibrium E_0 of the model is locally asymptotically stable; if $R_0>1$, it is unstable.

3. Optimal Control Strategy

Three bounded time-dependent control functions $u_1(t)$, $u_2(t)$ and $u_3(t)$ are introduced into the model. Here, $u_1(t)$ suppresses disease transmission by reducing the effective incidence rate, for example through mask wearing and avoidance of crowded places. The control $u_2(t)$ reduces the effective saturation pressure on medical resources, for example through hierarchical treatment, flexible scheduling, and simplified admission procedures. The control $u_3(t)$ represents treatment-efficiency enhancement, such as increasing medical staffing, optimizing treatment protocols, and improving the recovery rate of hospitalized individuals. The admissible control set is assumed to be

$$U=\Big\{(u_1,u_2,u_3)\,|\,u_i(t)\in L^\infty(0,T)^3,\\ 0\le u_i(t)\le u_{i,\max}, i=1,2,3\Big\},$$

where $u_{i,\max}$ represents the maximum achievable intensity of each intervention strategy.

The optimal control model is established as follows:

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = \left(\dfrac{A}{N}+b\right)N - \dfrac{(1-u_1)\beta SI}{\alpha I^2+1} - \mu S \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \dfrac{(1-u_1)\beta SI}{\alpha I^2+1} - rI\left(1-\dfrac{(1-u_2)H}{k}\right) - v_1 I - \varepsilon_1 I - \mu I \\ \dfrac{\mathrm{d}H}{\mathrm{d}t} = rI\left(1-\dfrac{(1-u_2)H}{k}\right) - (v_2+u_3)H - \varepsilon_2 H - \mu H \\ \dfrac{\mathrm{d}R}{\mathrm{d}t} = v_1 I + (v_2+u_3)H - \mu R \end{cases} . \tag{2}$$

In the proposed model, the term $\dfrac{H}{k}$ represents the original medical-resource saturation level. The introduction of $u_2(t)$ reduces the effective saturation level to $\dfrac{(1-u_2(t))H}{k}$, which reflects the improvement in healthcare efficiency.

3.1. Existence of an Optimal Control

To reduce the numbers of infected and hospitalized individuals while controlling intervention costs, we define the objective functional as

$$J(u_1,u_2,u_3) = \int_0^T \left(A_1 I + A_2 H + \frac{B_1}{2}u_1^2 + \frac{B_2}{2}u_2^2 + \frac{B_3}{2}u_3^2\right)\mathrm{d}t \tag{3}$$

where T is the terminal time; $A_1, A_2 > 0$ represent the social costs associated with infected and hospitalized individuals, respectively; and $B_1, B_2, B_3 > 0$ represent the implementation costs of the three control measures. These coefficients are selected to balance epidemic reduction and intervention costs in the numerical experiments.

Theorem 1: The optimal control problem defined by system (2) and objective functional (3) admits an optimal control

$$u^*(t) = (u_1^*, u_2^*, u_3^*) \in U$$

such that

$$J(u_1^*, u_2^*, u_3^*) = \min\{J(u_1,u_2,u_3) \mid (u_1,u_2,u_3) \in U\}$$

Proof: To prove the existence of the optimal control, we verify the conditions of the Fleming-Rishel optimal control existence theorem.

First, the admissible control set U is nonempty, closed, and convex.

Second, for any $u_i \in U, i = 1,2,3$, since $1+\alpha I^2 > 0$, the right-hand side of system (2) is continuous with respect to the state variables in the feasible region. Moreover, the system satisfies the local Lipschitz condition with respect to the state variables. Therefore, the state system admits a unique solution for every admissible control.

Third, since the uncontrolled system (1) is bounded, the controlled system (2) is also bounded. In addition, the terms related to the control variables in system (2) can be rewritten as

$$\frac{(1-u_1)\beta SI}{\alpha I^2+1}=\frac{\beta SI}{\alpha I^2+1}-\frac{u_1\beta SI}{\alpha I^2+1},$$
$$rI\left(1-\frac{(1-u_2)H}{k}\right)=rI-\frac{rIH}{k}+\frac{u_2 rIH}{k},$$

and

$$(v_2+u_3)H=v_2H+u_3H .$$

Hence, the state system is linear with respect to the control variables u_1,u_2,u_3.

Fourth, the integrand of the objective functional is convex with respect to the control variables. Define

$$L(I,H,u_1,u_2,u_3)=A_1I+A_2H+\frac{B_1}{2}u_1^2+\frac{B_2}{2}u_2^2+\frac{B_3}{2}u_3^2.$$

Since the quadratic terms are positive and $B_1,B_2,B_3>0$, L is convex with respect to $u_1(t)$, $u_2(t)$, $u_3(t)$.

Finally, the integrand satisfies the required growth condition:

$$L(I,H,u_1,u_2,u_3)\geq\frac{1}{2}\min\{B_1,B_2,B_3\}\left(u_1^2+u_2^2+u_3^2\right).$$

Therefore, all the assumptions of the Fleming-Rishel optimal control existence theorem are satisfied. Hence, there exists an optimal control

$$u^*(t)=\left(u_1^*,u_2^*,u_3^*\right)\in U,$$

such that

$$J\left(u_1^*,u_2^*,u_3^*\right)=\min\left\{J(u_1,u_2,u_3)\,|\,(u_1,u_2,u_3)\in U\right\}.$$

3.2. Characterization of Optimal Controls

The optimal control problem can be transformed into an equivalent Hamiltonian system by applying Pontryagin's Maximum Principle [8]. For this purpose, the Hamiltonian function is defined as

$$\begin{aligned}
H=&A_1I+A_2H+\frac{B_1}{2}u_1^2+\frac{B_2}{2}u_2^2+\frac{B_3}{2}u_3^2\\
&+\lambda_1\left[\left(\frac{A}{N}+b\right)N-\frac{(1-u_1)\beta SI}{\alpha I^2+1}-\mu S\right]\\
&+\lambda_2\left[\frac{(1-u_1)\beta SI}{\alpha I^2+1}-rI\left(1-\frac{(1-u_2)H}{k}\right)-v_1I-\varepsilon_1I-\mu I\right]\\
&+\lambda_3\left[rI\left(1-\frac{(1-u_2)H}{k}\right)-(v_2+u_3)H-\varepsilon_2H-\mu H\right]\\
&+\lambda_4\left[v_1I+(v_2+u_3)H-\mu R\right]
\end{aligned}$$

where $\lambda_i, i=1,2,3,4$ are the adjoint variables associated with the state variables S,I,H,R, respectively.

Theorem 2: Let $u^*=\left(u_1^*,u_2^*,u_3^*\right)$ be the optimal control. Then there exist adjoint variables $\lambda_i, i=1,2,3,4$ satisfying the following adjoint system:

$$\begin{cases} \dfrac{\mathrm{d}\lambda_1}{\mathrm{d}t} = -\left[\lambda_1\left(b-\mu-\dfrac{(1-u_1)\beta I}{1+\alpha I^2}\right)+\lambda_2\dfrac{(1-u_1)\beta I}{1+\alpha I^2}\right] \\ \dfrac{\mathrm{d}\lambda_2}{\mathrm{d}t} = -\left[A_1+\lambda_1\left(b-\dfrac{(1-u_1)\beta S(1-\alpha I^2)}{(1+\alpha I^2)^2}\right)\right. \\ \qquad +\lambda_2\left(\dfrac{(1-u_1)\beta S(1-\alpha I^2)}{(1+\alpha I^2)^2}-r\left(1-\dfrac{(1-u_2)H}{k}\right)-(v_1+\varepsilon_1+\mu)\right) \\ \qquad \left.+\lambda_3 r\left(1-\dfrac{(1-u_2)H}{k}\right)+\lambda_4 v_1\right] \\ \dfrac{\mathrm{d}\lambda_3}{\mathrm{d}t} = -\left[A_2+b\lambda_1+\lambda_2\dfrac{r(1-u_2)I}{k}+\lambda_3\left(-\dfrac{r(1-u_2)I}{k}-(v_2+\varepsilon_2+\mu)-u_3\right)+\lambda_4(v_2+u_3)\right] \\ \dfrac{\mathrm{d}\lambda_4}{\mathrm{d}t} = -\left[b\lambda_1-\mu\lambda_4\right] \end{cases}, \quad (4)$$

with the transversality conditions $\lambda_i(T)=0, i=1,2,3,4$. The optimal controls are given by

$$u_1^* = \min\left\{\max\left\{0,\frac{(\lambda_2-\lambda_1)\beta SI}{B_1(1+\alpha I^2)}\right\},u_{1,\max}\right\}, \quad (5)$$

$$u_2^* = \min\left\{\max\left\{0,\frac{(\lambda_2-\lambda_3)rIH}{B_2 k}\right\},u_{2,\max}\right\}, \quad (6)$$

$$u_3^* = \min\left\{\max\left\{0,\frac{(\lambda_3-\lambda_4)H}{B_3}\right\},u_{3,\max}\right\}. \quad (7)$$

Proof: According to Pontryagin's Maximum Principle, there exist adjoint variables $\lambda_i, i=1,2,3,4$ such that the Hamiltonian satisfies $\dfrac{\mathrm{d}\lambda_i}{\mathrm{d}t}=-\dfrac{\partial H}{\partial x_i}, x_i=S,I,H,R$, with the transversality conditions $\lambda_i(T)=0, i=1,2,3,4$.

Taking the partial derivatives of the Hamiltonian with respect to the state variables S,I,H,R, respectively, yields the adjoint system (5).

Furthermore, the optimality conditions satisfy $\dfrac{\mathrm{d}H}{\mathrm{d}u_i}=0, i=1,2,3$.

Solving these equations gives the corresponding interior optimal controls:

$$\tilde{u}_1=\frac{(\lambda_2-\lambda_1)\beta SI}{B_1(1+\alpha I^2)},$$

$$\tilde{u}_2=\frac{(\lambda_2-\lambda_3)rIH}{B_2 k},$$

$$\tilde{u}_3=\frac{(\lambda_3-\lambda_4)H}{B_3}.$$

Considering the control constraints

$$0 \le u_i(t) \le u_{i,\max}, i = 1,2,3,$$

the optimal controls are obtained by projecting the interior solutions onto the admissible control set, which leads to Equations (6)-(8).

4. Numerical Simulations

In this section, numerical simulations are performed to verify the proposed optimal control strategy and to investigate the effects of different control measures on the number of infected individuals, hospitalized individuals, and medical-resource saturation pressure.

The numerical simulations are conducted using MATLAB. The state system is solved by the fourth-order Runge-Kutta method, while the adjoint system is solved backward in time according to the terminal conditions. The optimal control variables are updated by using the forward-backward sweep method [8].

The simulation time interval is set as $[0,160]$ and the initial values are given as follows:

$$\left(S(0), I(0), H(0), R(0)\right) = (150, 20, 6, 20).$$

4.1. Parameter Values and Verification of Theoretical Conditions

The parameter values used in the numerical simulations are selected based on biological plausibility and previous epidemic modeling studies. Since this work focuses on evaluating the qualitative effects of different control strategies, the parameters are not obtained through calibration with a specific outbreak dataset. The parameter values used in the numerical simulations are selected as follows (**Table 1**):

Table 1. Parameter values used in the numerical simulations.

Parameter	value	Parameter	value
A	16.4947	b	0.0441
μ	0.1064	β	0.0074
α	0.0002	r	0.2066
k	16.8749	v_1	0.1474
v_2	0.1924	ε_1	0.0084
ε_2	0.0122		

For these parameter values, the basic reproduction number is calculated as $R_0 = 4.183 > 1$.

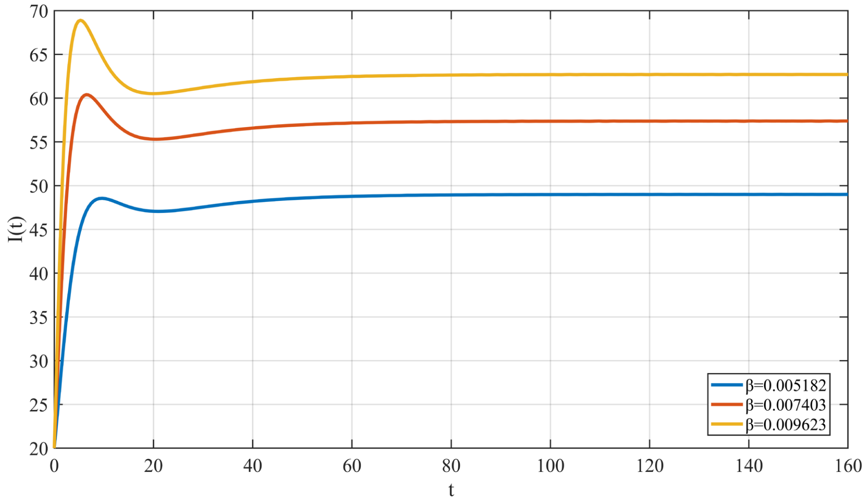
The control cost parameters are selected as $A_1 = 1$, $A_2 = 1.4$, $B_1 = 120$, $B_2 = 8$, $B_3 = 80$.

Considering the limitations of medical technology and implementation costs, the upper bounds of the three control variables are set as $u_{1,\max} = 0.75$, $u_{2,\max} = 0.95$, $u_{3,\max} = 0.35$.

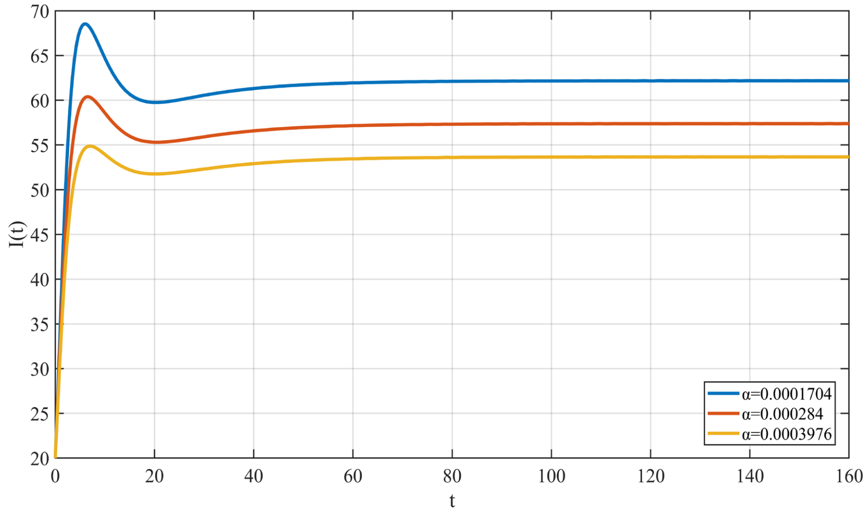
4.2. Effects of Key Parameters on Infected and Hospitalized Individuals

To investigate the effects of the main parameters on disease transmission dynamics, the parameters β, α, k, and r are varied individually while keeping the other parameters unchanged under the uncontrolled scenario. Figure 1 and Figure 2 illustrate the effects of these parameters on the number of infected individuals $I(t)$ and hospitalized individuals $H(t)$, respectively.

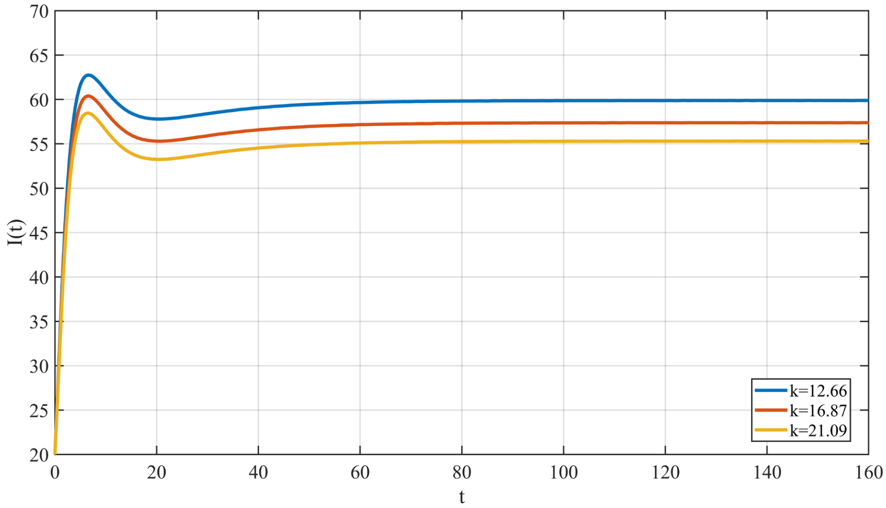
As shown in Figure 1, the number of infected individuals increases significantly with the increase of the transmission rate β, indicating that reducing the effective transmission rate is a key measure for controlling the scale of disease transmission. With the increase of the nonlinear inhibition coefficient α, the number of infected individuals decreases due to the enhanced inhibitory effect of the nonlinear incidence term. This result reflects the potential role of preventive awareness and social intervention measures in reducing disease transmission.



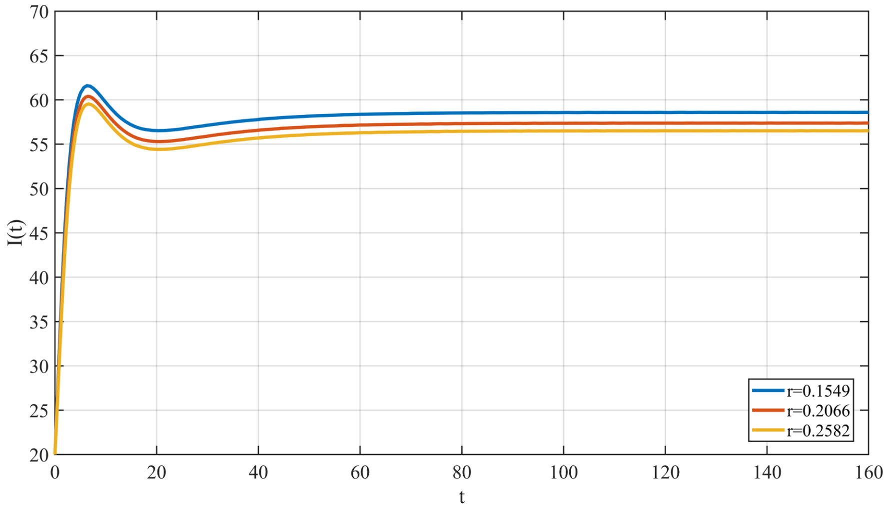
(a)



(b)

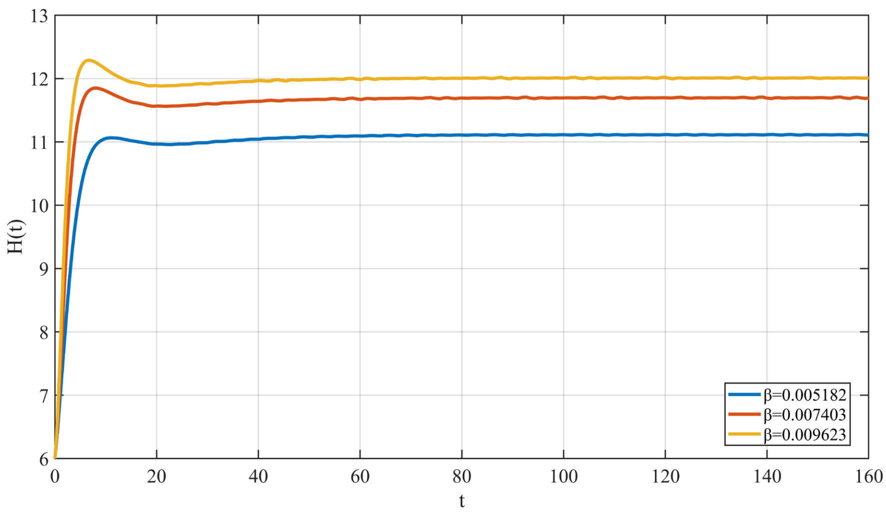


(c)

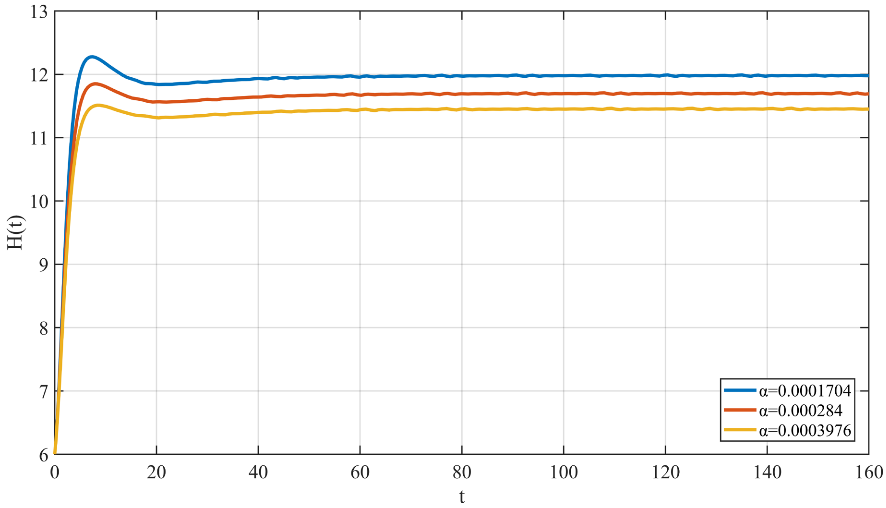


(d)

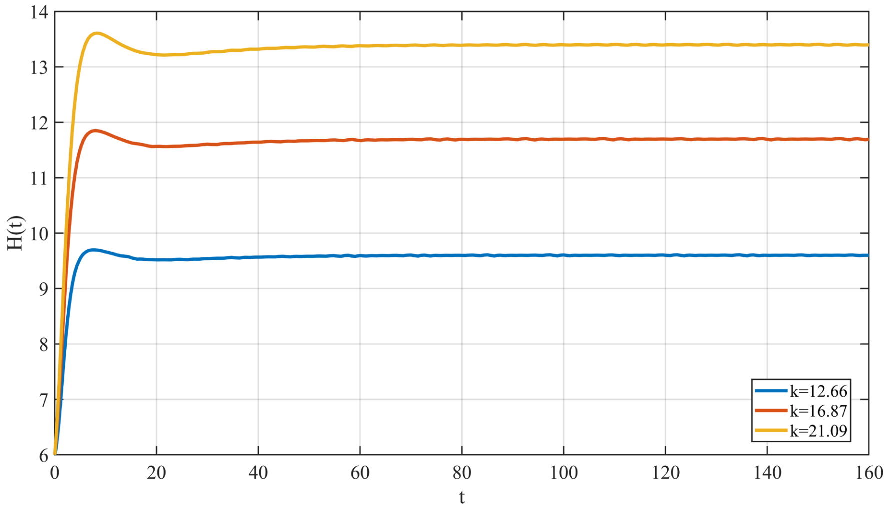
Figure 1. Effects of β, α, k, and r on $I(t)$.



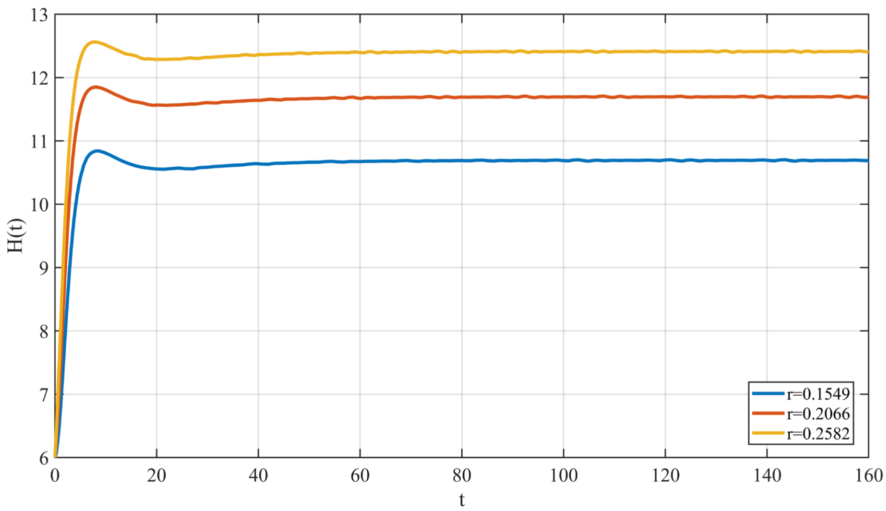
(a)



(b)



(c)



(d)

Figure 2. Effects of β, α, k, and r on $H(t)$.

Comparing **Figure 1** and **Figure 2**, it can be observed that the medical capacity k and the hospitalization transfer rate r jointly affect the distribution relationship between infected individuals and hospitalized individuals. A larger k alleviates the limitation of hospital capacity, allowing more infected individuals to receive hospitalization treatment. Consequently, the number of infected individuals $I(t)$ decreases, while the number of hospitalized individuals $H(t)$ increases. A larger r accelerates the transition from infected individuals to hospitalized individuals, which is reflected by a reduction in the number of infected individuals and an increase in the number of hospitalized individuals. These results indicate that medical capacity and hospitalization transfer rate do not directly eliminate infections, but rather modify the dynamic transition process between infected and hospitalized compartments.

4.3. Comparison of Different Control Strategies

To evaluate the effects of the three control measures, the following five control strategies are considered:

1) No control, namely $u_1 = u_2 = u_3 = 0$;

2) Only the transmission reduction control u_1;

3) Only the medical-resource saturation mitigation control u_2;

4) Only the treatment efficiency improvement control u_3;

5) The combined control strategy involving u_1, u_2, u_3.

As shown in **Figure 3** and **Figure 4**, different control strategies have distinct effects on the dynamics of infected and hospitalized individuals. When only the transmission reduction control u_1 is implemented, the number of infected individuals decreases significantly, and the number of hospitalized individuals also decreases accordingly. This indicates that reducing the effective transmission rate plays an important role in controlling disease spread.

When only the treatment efficiency improvement control u_3 is applied, the number of hospitalized individuals is significantly reduced, while its effect on reducing the number of infected individuals is relatively limited. This is because u_3 mainly improves the recovery efficiency of hospitalized individuals rather than directly affecting the infection process.

When only the medical-resource saturation mitigation control u_2 is adopted, the number of infected individuals is close to that in the uncontrolled scenario, and the number of hospitalized individuals does not decrease significantly, or may even be slightly higher than that without control. This is because the primary role of u_2 is not to directly reduce infections or accelerate recovery, but to alleviate the restriction caused by limited medical capacity on the hospitalization process.

The combined control strategy simultaneously reduces transmission risk, alleviates medical-resource saturation, and improves treatment efficiency. Therefore, it achieves the most significant reduction in both $I(t)$ and $H(t)$, demonstrating the superiority of the integrated intervention strategy.

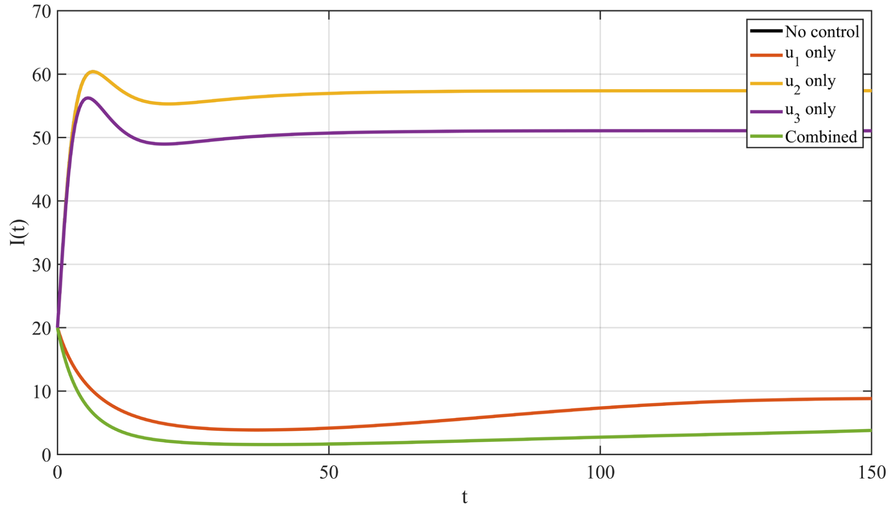


Figure 3. Dynamics of $I(t)$ under different control strategies.

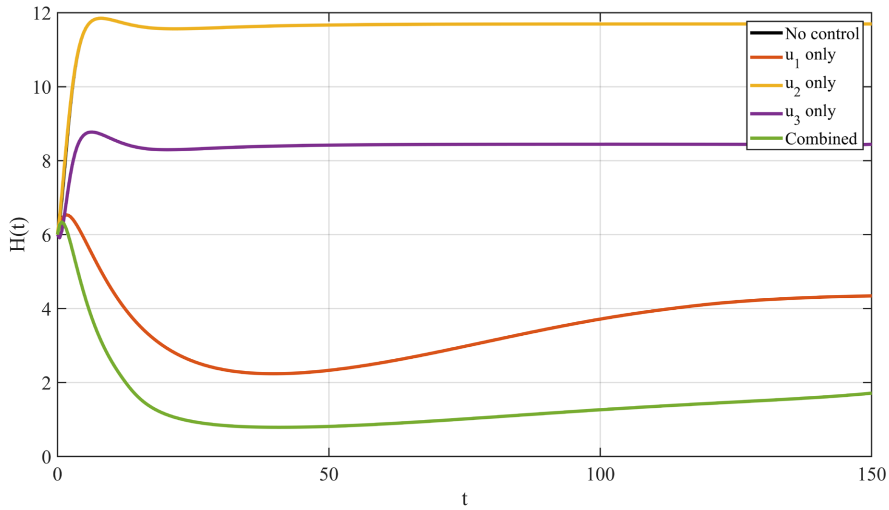


Figure 4. Dynamics of $H(t)$ under different control strategies.

4.4. Optimal Control Profiles under the Combined Control Strategy

Figure 5 presents the optimal trajectories of the three control variables under the combined control strategy. It can be observed that the transmission reduction control u_1 remains almost at its upper bound throughout the entire simulation period. This indicates that, under the selected weighting parameters, reducing the effective transmission rate is the dominant intervention strategy for controlling disease transmission.

The medical-resource saturation mitigation control u_2 initially remains close to its maximum value during the early stage of the epidemic, indicating that medical-resource allocation is particularly important when the disease burden is high. Subsequently, u_2 decreases rapidly and remains at a moderate level during the later period, suggesting that the pressure on healthcare resources gradually declines as the number of infected individuals decreases.

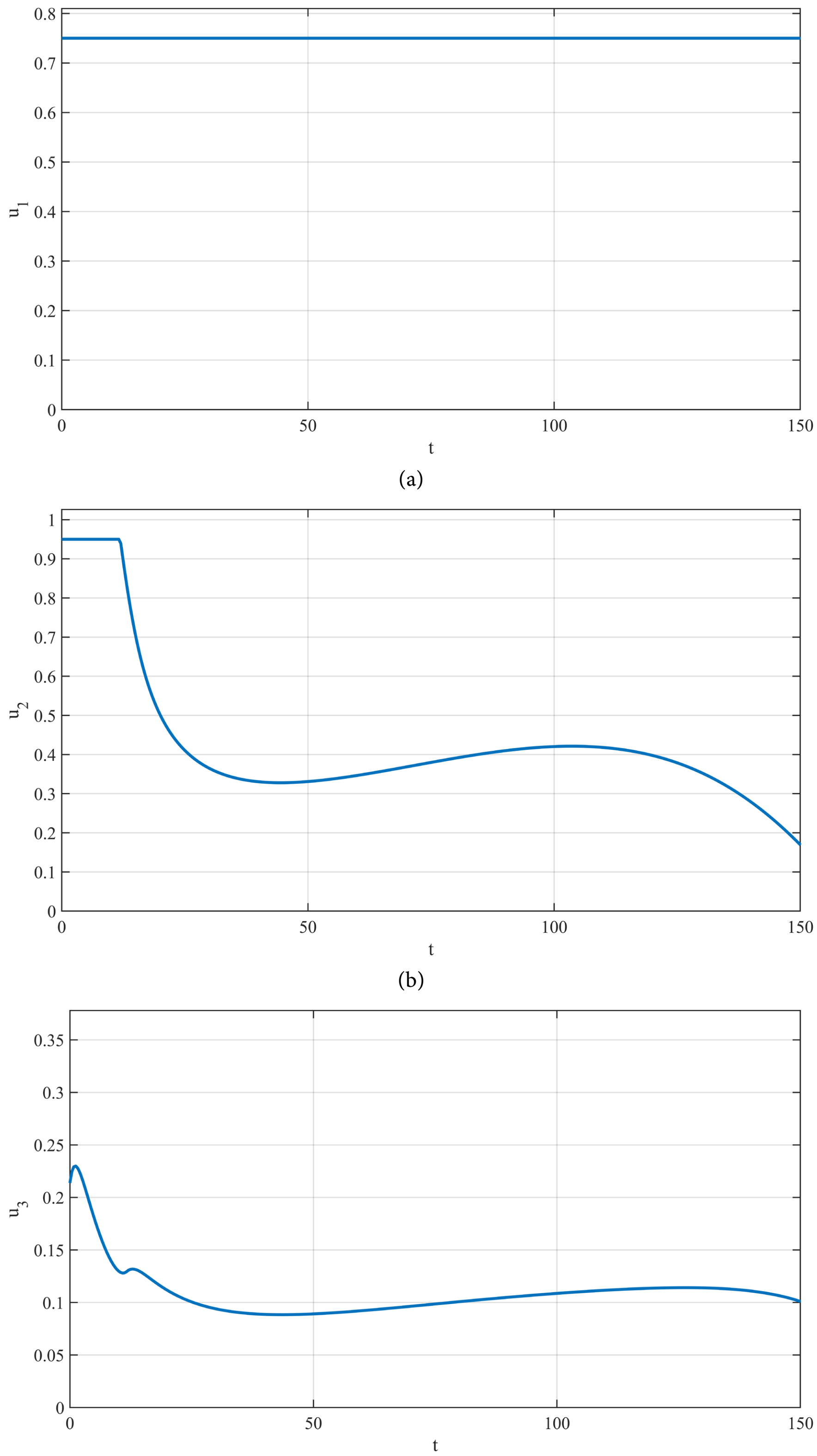


Figure 5. Optimal trajectories of u_1, u_2, u_3 under the combined control strategy.

The treatment efficiency improvement control u_3 maintains a positive value throughout the simulation period, although its magnitude is lower than those of u_1 and u_2. This indicates that improving treatment efficiency serves as a continuous auxiliary measure, which works together with transmission reduction and medical-resource management to alleviate hospitalization pressure.

These results demonstrate that, under the given parameter settings and cost weights, reducing the effective transmission rate remains the most cost-effective strategy for controlling disease spread, while medical-resource allocation and treatment enhancement play important complementary roles.

4.5. Comparison of Cumulative Disease Burden under Different Control Strategies

Since the social cost terms in the objective functional are positively associated with $I(t)$ and $H(t)$, the cumulative disease burden can be used to evaluate the long-term effects of different control strategies. Therefore, we define $B_I = \int_0^T I(t)\mathrm{d}t$ and $B_H = \int_0^T H(t)\mathrm{d}t$, which represent the cumulative infection burden and cumulative hospitalization burden, respectively.

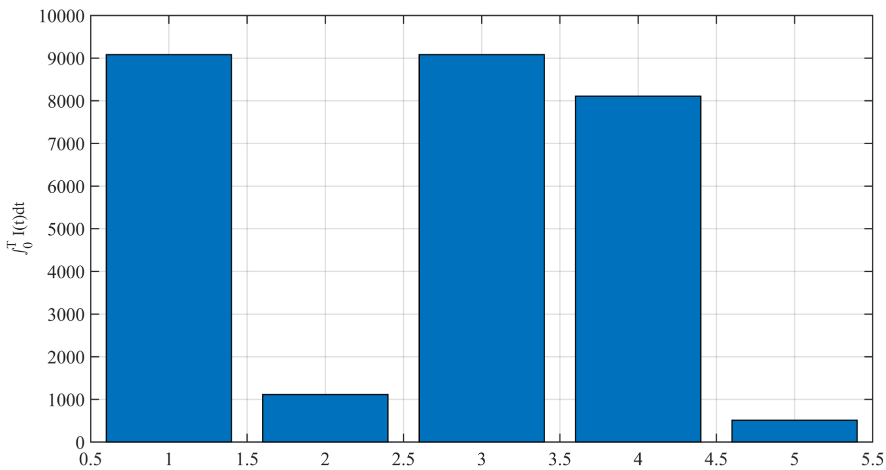


Figure 6. Cumulative infection burden under different control strategies.

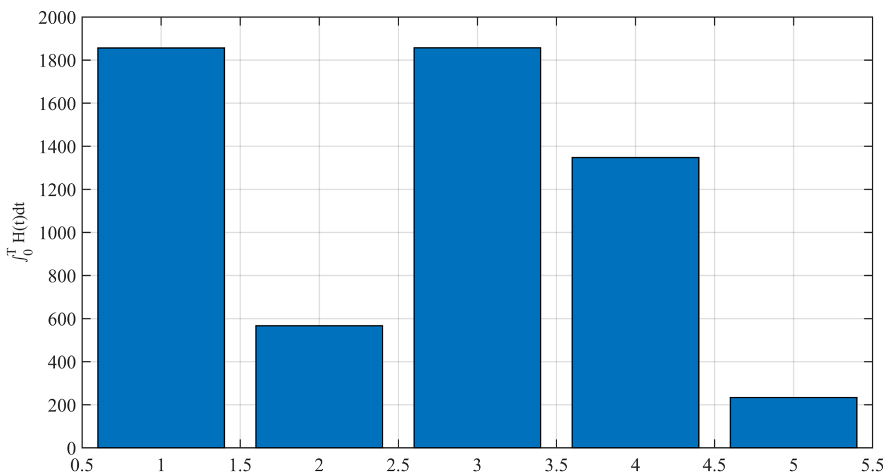


Figure 7. Cumulative hospitalization burden under different control strategies.

Figure 6 and Figure 7 illustrate the comparison of these two indicators under different control strategies. It can be observed that the combined control strategy achieves the lowest cumulative infection burden and cumulative hospitalization burden.

Specifically, under the no-control scenario, the cumulative infection burden and cumulative hospitalization burden are $B_I = 9080.7874$ and $B_H = 1856.3709$, respectively. Under the combined control strategy, these values decrease to 512.8115 and 233.6586, respectively.

When only the transmission reduction control u_1 is implemented, the cumulative infection burden and cumulative hospitalization burden decrease to 1116.5956 and 566.7082, respectively, indicating that reducing the effective transmission rate plays a significant role in controlling disease spread.

When only the treatment efficiency improvement control u_3 is applied, the cumulative hospitalization burden decreases to 1347.2779, while the cumulative infection burden remains relatively high at 8109.1681. This demonstrates that improving treatment efficiency mainly alleviates hospitalization pressure but has a limited effect on reducing infection transmission.

When only the medical-resource saturation mitigation control u_2 is adopted, the cumulative infection burden is 9080.3431, which is close to the uncontrolled scenario, while the cumulative hospitalization burden is 1856.6231, slightly higher than that without control. This result indicates that medical-resource saturation mitigation mainly improves the capacity for hospitalization rather than directly reducing the number of infections. Therefore, it should be combined with transmission reduction and treatment enhancement strategies to achieve better epidemic control.

4.6. Analysis of Medical Resource Saturation Pressure

Since u_2 is introduced to alleviate medical-resource saturation, the effective saturation pressure $\frac{(1-u_2)H(t)}{k}$ is further investigated.

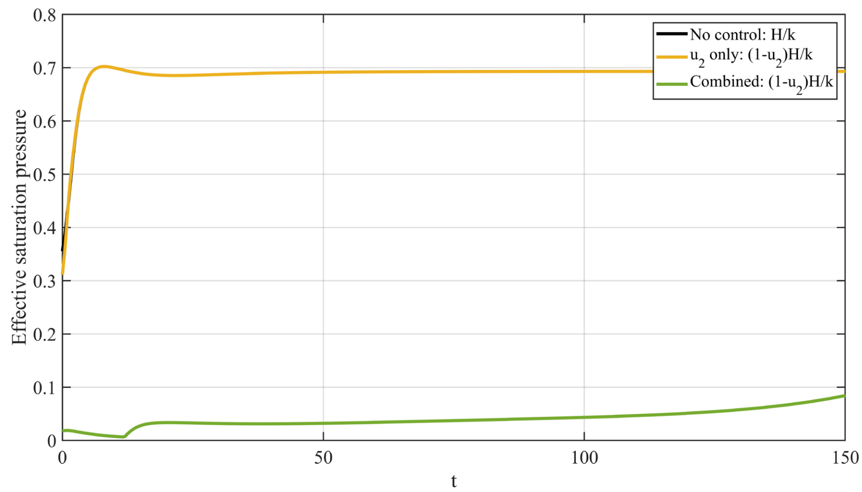


Figure 8. Medical-resource saturation pressure under three different control strategies.

As shown in **Figure 8**, without control, the medical-resource saturation pressure $\frac{H}{k}$ remains at a relatively high level throughout the simulation period. When only the medical-resource saturation mitigation control u_2 is applied, this indicator does not decrease significantly, indicating that increasing medical-resource allocation alone cannot fundamentally reduce hospitalization pressure.

In contrast, the combined control strategy significantly reduces the effective saturation pressure. This result demonstrates that medical-resource mitigation measures need to be implemented together with transmission reduction and treatment efficiency improvement strategies to achieve better epidemic control effects.

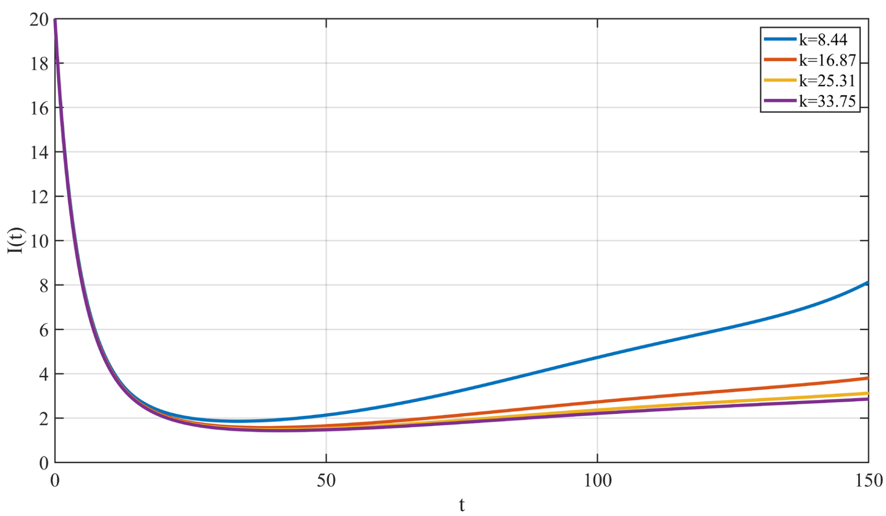
4.7. Evaluation of the Effect of Medical Capacity Expansion on Epidemic Control

The medical capacity parameter k represents a key indicator of limited healthcare resources in the proposed model. In system (1), the effective transition rate from infected individuals to hospitalized individuals is given by $rI\left(1-\frac{H}{k}\right)$, where $\frac{H}{k}$ represents the proportion of hospitalized individuals relative to the available medical capacity. When H approaches k, the ability of infected individuals to enter the hospitalization stage is significantly restricted due to healthcare resource saturation.

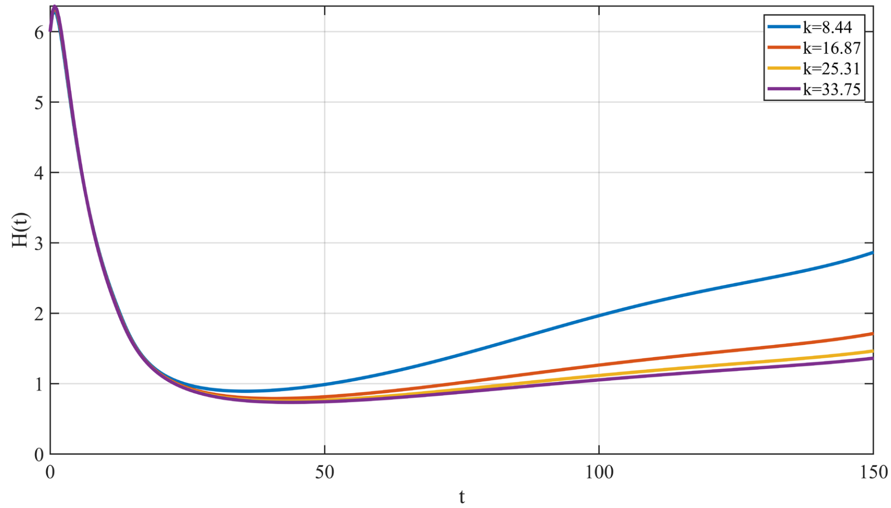
Therefore, in this section, the effects of different medical capacities on disease transmission dynamics are investigated while keeping all other parameters unchanged. The following four scenarios are considered:

$$k = 0.5k_0,\quad k = k_0,\quad k = 1.5k_0,\quad k = 2k_0.$$

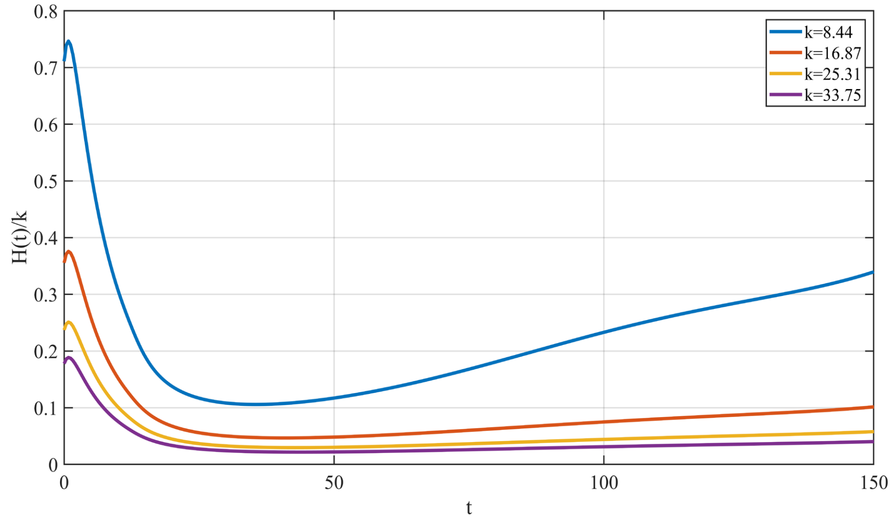
where $k_0 = 16.8749$. To highlight the role of medical capacity, the effects of different values of on the numbers of infected individuals, hospitalized individuals, and medical-resource saturation pressure are simulated under the combined optimal control strategy.



(a)



(b)



(c)

Figure 9. Effects of different medical capacities on $I(t)$, $H(t)$ and $\frac{H}{k}$.

As shown in **Figure 9**, the number of infected individuals decreases significantly with increasing medical capacity k. This is because an increase in healthcare capacity weakens the restriction imposed by limited medical resources on the hospitalization transition process, allowing infected individuals to enter the hospitalization stage more efficiently and reducing the number of individuals who continue transmitting the disease in the community.

In **Figure 9(b)**, the hospitalization curves under different values of k exhibit similar peak values during the early stage of the epidemic. However, smaller medical capacities lead to higher numbers of hospitalized individuals in the later stage, indicating that insufficient healthcare capacity prolongs hospitalization pressure.

More importantly, the medical-resource saturation pressure $\frac{H}{k}$ decreases significantly as k increases, demonstrating that medical capacity expansion can effectively alleviate healthcare system saturation.

Furthermore, the peak number of infected individuals $I_{\max}$, the peak number of hospitalized individuals $H_{\max}$, the cumulative infection burden $B_I = \int_0^T I(t)\mathrm{d}t$, the cumulative hospitalization burden $B_H = \int_0^T H(t)\mathrm{d}t$, and the maximum medical-resource saturation pressure $\max\left(\frac{H}{k}\right)$ are calculated. The results are presented in Table 2.

Table 2. Effects of different medical capacities on epidemic indicators.

k	$I_{\max}$	$H_{\max}$	B_I	B_H	$\max\left(\frac{H}{k}\right)$
8.4375	47.4327	6.3003	796.1781	315.9099	0.7467
16.8749	32.7209	6.3418	512.8115	233.6586	0.3758
25.3124	28.6782	6.3556	460.8530	215.4125	0.2511
33.7498	26.9087	6.3626	439.4709	207.5966	0.1885

As shown in Table 2, when the medical capacity k increases from 8.4375 to 33.7498, the peak number of infected individuals decreases from 47.4327 to 26.9087. Meanwhile, the cumulative infection burden decreases from 796.1782 to 439.4709, and the cumulative hospitalization burden decreases from 315.9099 to 207.5966.

At the same time, the maximum medical-resource saturation pressure decreases significantly from 0.7467 to 0.1885. Although the peak number of hospitalized individuals $H_{\max}$ slightly increases with the increase of k, this phenomenon occurs because a larger medical capacity weakens the restriction caused by limited healthcare resources, allowing more infected individuals to enter the hospitalization stage during the early period of the epidemic. From the perspectives of the cumulative hospitalization burden and the saturation pressure $\frac{H}{k}$, a larger medical capacity can effectively alleviate long-term hospitalization pressure.

Therefore, the expansion of medical capacity not only reduces the saturation pressure of the healthcare system but also decreases the peak number of infected individuals and the overall disease burden by improving the transition efficiency from infected individuals to hospitalized individuals. This result highlights the importance of incorporating medical capacity constraints into epidemic models. In practical epidemic control, increasing healthcare capacity should not only be regarded as improving hospital admission capability, but also as an intervention measure that can influence disease dynamics by modifying the transition process between community infection and hospitalization.

In summary, the numerical results demonstrate that although individual control measures can reduce the number of infected or hospitalized individuals to some extent, their effects remain limited. The combined control strategy, together with sufficient medical capacity, can simultaneously reduce infection burden, hospitalization burden, and medical-resource saturation pressure, providing a more effective and comprehensive approach for epidemic control.

5. Conclusions

In this paper, an optimal control problem for an SIHR epidemic model with nonlinear birth rate, nonlinear incidence rate, and limited medical resources are investigated. By introducing three time-dependent control variables, including transmission reduction control, medical-resource saturation mitigation control, and treatment efficiency improvement control, an optimal control model is established. The existence of optimal controls is proved by applying Fleming-Rishel's optimal control existence theorem, and the necessary conditions for the optimal controls are derived based on Pontryagin's Maximum Principle.

The numerical results demonstrate that the combined control strategy can simultaneously reduce the number of infected individuals, hospitalized individuals, and the overall disease burden, and its performance is superior to that of any single control strategy. Among the three control measures, transmission reduction plays the most important role in suppressing disease transmission, treatment efficiency improvement is effective in alleviating hospitalization pressure, and medical-resource saturation mitigation mainly influences the transition process from infection to hospitalization.

Furthermore, the results indicate that increasing medical capacity can effectively alleviate healthcare-system saturation pressure and reduce disease transmission by improving the hospitalization transition process. These findings suggest that coordinated intervention strategies and rational allocation of medical resources are of great significance for epidemic prevention and control. Future research may extend the proposed model by incorporating time delays, stochastic effects, and vaccination strategies to further improve its applicability in describing realistic epidemic transmission dynamics.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

[1] Kermack, W.O. and McKendrick, A.G. (1927) A Contribution to the Mathematical Theory of Epidemics. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, **115**, 700-721. https://doi.org/10.1098/rspa.1927.0118

[2] Ruan, S. and Wang, W. (2003) Dynamical Behavior of an Epidemic Model with a Nonlinear Incidence Rate. *Journal of Differential Equations*, **188**, 135-163. https://doi.org/10.1016/s0022-0396(02)00089-x

[3] Zhang, J., Jia, J. and Song, X. (2014) Analysis of an SEIR Epidemic Model with Saturated Incidence and Saturated Treatment Function. *The Scientific World Journal*, **2014**, Article ID: 910421. https://doi.org/10.1155/2014/910421

[4] Chen, C. and Xiang, Z. (2025) Analysis and Optimal Control Strategies of Epidemic

Models with Nonlinear Incidence Rates. *Advances in Continuous and Discrete Models*, **2025**, Article No. 92. https://doi.org/10.1186/s13662-025-03950-8

[5] Zhang, X. and Liu, X. (2008) Backward Bifurcation of an Epidemic Model with Saturated Treatment Function. *Journal of Mathematical Analysis and Applications*, **348**, 433-443. https://doi.org/10.1016/j.jmaa.2008.07.042

[6] Shan, C. and Zhu, H. (2014) Bifurcations and Complex Dynamics of an SIR Model with the Impact of the Number of Hospital Beds. *Journal of Differential Equations*, **257**, 1662-1688. https://doi.org/10.1016/j.jde.2014.05.030

[7] Wang, F., Liu, M., Zhang, L. and Xie, B. (2023) Bifurcation Analysis and Optimal Control of a Network-Based SIR Model with the Impact of Medical Resources. *Nonlinear Analysis: Modelling and Control*, **28**, 1-19. https://doi.org/10.15388/namc.2023.28.30770

[8] Laarabi, H., Labriji, E.H.H., Rachik, M. and Kaddar, A. (2012) Optimal Control of an Epidemic Model with a Saturated Incidence Rate. *Nonlinear Analysis: Modelling and Control*, **17**, 448-459. https://doi.org/10.15388/na.17.4.14050

[9] Lenhart, S. and Workman, J.T. (2007) Optimal Control Applied to Biological Models. Chapman & Hall & CRC.

[10] Balderrama, R., Peressutti, J., Pinasco, J.P., Vazquez, F. and Vega, C.S.D.L. (2022) Optimal Control for a SIR Epidemic Model with Limited Quarantine. *Scientific Reports*, **12**, Article No. 12583. https://doi.org/10.1038/s41598-022-16619-z

[11] Du, J., Qin, C. and Hui, Y. (2024) Optimal Control and Analysis of a Stochastic SEIR Epidemic Model with Nonlinear Incidence and Treatment. *AIMS Mathematics*, **9**, 33532-33550. https://doi.org/10.3934/math.20241600