



Stochastic Modeling of Epidemic Extinction Under Non-pharmaceutical Public-health Policy Interventions

Lixia Wang*, and Donald Porchia

Department of Biostatistics, University of Florida, Gainesville, FL 32610, United States.

Article Details

Article Type: Research Article

Received date: 03rd November, 2025

Accepted date: 20th February, 2026

Published date: 24th February, 2026

***Corresponding Author:** Lixia Wang, Clinical Assistant Professor, Department of Biostatistics, University of Florida, Gainesville, FL 32610, United States.

Citation: Wang, L., & Porchia, D., (2026). Stochastic Modeling of Epidemic Extinction Under Non-pharmaceutical Public-health Policy Interventions. *J Pub Health Issue Pract* 10(1): 253. doi: <https://doi.org/10.33790/jphip1100253>

Copyright: ©2026, This is an open-access article distributed under the terms of the [Creative Commons Attribution License 4.0](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.

Abstract

Public health interventions can change epidemic trajectories and, critically, can also increase the chance that an outbreak dies out early. We developed a quantitative framework to estimate the probability of epidemic extinction under different strengths of non-pharmaceutical interventions implemented in two subpopulations (low-risk and high-risk). The approach extends a standard compartmental model by incorporating random transmission events and evaluating disease extinction probabilities under alternative intervention scenarios.

We applied the framework to COVID-19 incidence data from Brazil (Johns Hopkins University). Under fitted baseline parameters, the basic reproduction number was approximately $R_0 = 1.4$, indicating that deterministic models would predict sustained epidemic growth. Despite $R_0 > 1$, the stochastic framework yields substantial extinction probabilities that can be increased through timely interventions. When the outbreak was seeded by one infectious individual in the high-risk group, the estimated extinction probability under baseline intervention levels was 67.22%. A 40% proportional increase in low-risk interventions (holding high-risk interventions at baseline) increased extinction to 74.93%, whereas the same proportional increase applied only to high-risk interventions increased extinction modestly to 68.43%. Increasing interventions in both groups by 40% yielded a similar extinction probability (75.13%), highlighting that broad population-level measures can achieve near-maximal gains relative to dual targeting at the same scale.

Overall, these results indicate that early, population-wide non-pharmaceutical interventions, particularly those improving compliance or reducing transmission in the larger low-risk group, can meaningfully increase the likelihood of epidemic control even when $R_0 > 1$.

Key words: Epidemic Extinction; Non-pharmaceutical Interventions; Compartmental Model; Markov Chain; Branching Process; COVID-19

Introduction

Public health interventions alter key epidemiological parameters, such as contact structure, transmission rate, and recovery rate, and

thereby influence the likelihood that an epidemic will naturally die out. Understanding the mechanisms that govern the probability of disease extinction is fundamental for assessing the effectiveness of non-pharmaceutical interventions. Although deterministic compartmental models have been widely used to study epidemic dynamics, the transmission of infectious diseases is inherently stochastic due to random contact patterns, demographic variability, and unobserved environmental effects. Incorporating stochasticity is therefore essential for capturing the range of possible epidemic outcomes, particularly the probability of stochastic extinction.

In this study, we extend the deterministic compartmental model previously described by Asempapa, R. et al [1] by developing a continuous-time Markov chain (CTMC) representation of the epidemic process and approximating it using a multitype branching process. This hybrid stochastic-deterministic framework enables us to derive the analytical characterization of the extinction probability as a function of intervention strengths. Using COVID-19 incidence data from Brazil (Johns Hopkins University), we conduct numerical simulations to illustrate stochastic trajectories of disease dynamics in different compartments. Extinction probabilities are estimated through simulation of the CTMC and computed directly from the branching-process formulation, providing a comprehensive assessment of how early-stage public health interventions shape the likelihood of epidemic extinction.

This study's objective was to quantify the probability that an emerging epidemic goes extinct under alternative public-health intervention strategies, with particular emphasis on differential targeting of low-risk and high-risk subpopulations. Extinction probability is a policy-relevant endpoint because it complements the basic reproduction number R_0 by translating transmission conditions into the practical likelihood that an outbreak will die out early, thereby informing when broad non-pharmaceutical measures are warranted and how aggressively they should be implemented. Using Brazil-parameterized COVID-19 dynamics, we show that proportional strengthening of interventions in the larger low-risk group produces substantially larger gains in extinction probability than the same proportional increases confined to the high-risk group,

and can achieve extinction probabilities comparable to strengthening interventions in both groups simultaneously.

Methods

Deterministic compartmental model

A compartmental model is a mathematical framework in which the population is divided into distinct epidemiological classes (compartments), with individuals in each class assumed to be homogeneous. Such models are typically represented as graphs linking the classes through transfer rates, which in turn define a system of differentia [1] equations governing disease dynamics [2]. Compartmental modeling dates back to W. H. Hamer [3] and has been developed extensively over more than a century. It is interesting to note that the foundations of this entire approach in modern epidemic modeling were created not by mathematicians but by public-health physicians, most prominently Hamer, Ross, Hamer, McKendrick, and Kermack, between 1900 and 1935 [4].

Asempapa et al. [1] proposed an eight- compartment model that stratifies the population into two subgroups: low-risk (including children and adults without underlying conditions) and high-risk (primarily elderly individuals or those with comorbidities). Within each subgroup, individuals progress through susceptible (S), exposed (E), and infectious (I) classes. Infectious individuals may subsequently transition to hospitalized (H), or recovered (R) states, yielding a total of eight epidemiological compartments. The model assumes a large, constant population size N , homogeneity within each compartment, infection rates proportional to the effective contact rate, and constant parameter values.

For clarity and ease of reference, we briefly summarize the structure of this model and the key results relevant to our analysis. After incorporating an underlying deceased (D) compartment, the resulting flowchart of the extended model is illustrated in Figure 1.

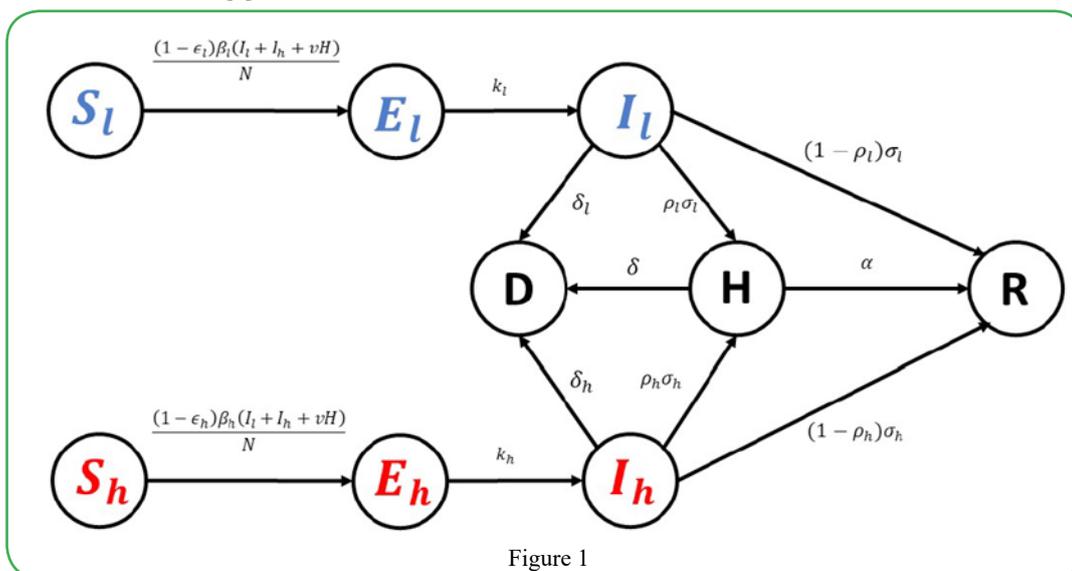


Figure 1

Figure 1. Deterministic compartmental model with low-risk and high-risk subpopulations. Flow diagram of the extended compartmental model stratified by risk group (low-risk: S_l, E_l, I_l ; high-risk: S_h, E_h, I_h) with shared hospitalized (H), recovered (R), and deceased (D) compartments. Arrows denote allowed transitions and model flows; parameters governing each transition are defined in Table 1.

Here, S_l, E_l , and I_l represent the Susceptible, Exposed, and Infectious groups, as well as their corresponding population sizes within the low-risk population, respectively. Similarly, S_h, E_h , and I_h denote Susceptible, Exposed, and Infectious groups and their associated population sizes in the high-risk subpopulation, respectively. The compartments D, H , and R refer to the Deceased, Hospitalized, and Recovered groups and their population sizes, respectively. The population sizes of all nine compartments vary over time.

The corresponding deterministic compartmental model is then described by the following system of differential equations, with parameter definitions summarized in Table 1:

$$\begin{aligned}
 S'_l(t) &= -\beta_l(1 - \epsilon_l)(I_l(t) + I_h(t) + vH(t)) \frac{S_l(t)}{N} \\
 E'_l(t) &= \beta_l(1 - \epsilon_l)(I_l(t) + I_h(t) + vH(t)) \frac{S_l(t)}{N} - k_l E_l(t) \\
 I'_l(t) &= k_l E_l(t) - (\sigma_l + \delta_l) I_l(t) \\
 S'_h(t) &= -\beta_h(1 - \epsilon_h)(I_l(t) + I_h(t) + vH(t)) \frac{S_h(t)}{N} \\
 E'_h(t) &= \beta_h(1 - \epsilon_h)(I_l(t) + I_h(t) + vH(t)) \frac{S_h(t)}{N} - k_h E_h(t) \\
 I'_h(t) &= k_h E_h(t) - (\sigma_h + \delta_h) I_h(t) \\
 H'(t) &= \rho_l \sigma_l I_l(t) + \rho_h \sigma_h I_h(t) - \alpha H(t) - \delta H(t) \\
 R'(t) &= (1 - \rho_l) \sigma_l I_l(t) + (1 - \rho_h) \sigma_h I_h(t) + \alpha H(t) \\
 D'(t) &= \delta_l I_l(t) + \delta_h I_h(t) + \delta H(t)
 \end{aligned}
 \tag{1}$$

The basic Reproduction Number R_0 , which is the average number of infectious individuals generated by one infected individual introduced into the entire susceptible population over the entire duration of the infectious period, is given by

$$R_0 = \frac{(1-\epsilon_l)\beta_l(\alpha+v\sigma_l\rho_l)S_l(0)}{\alpha(\sigma_l+\delta_l)N} + \frac{(1-\epsilon_h)\beta_h(\alpha+v\sigma_h\rho_h)S_h(0)}{\alpha(\sigma_h+\delta_h)N} = R_l + R_h \tag{2}$$

Here, $S_l(0)$ and $S_h(0)$ denote the initial sizes of the susceptible populations in the low-risk and high-risk subgroups, respectively, and $S_l(0) + S_h(0) = N$.

Table 1. Model parameters and definitions for the deterministic compartmental model. Definitions of parameters used in system (1) and Eq. (2), including transmission/contact rates, intervention parameters ϵ_l and ϵ_h , progression rates, hospitalization/recovery rates, and mortality rates. Units are day⁻¹ unless otherwise specified; dimensionless parameters (e.g., proportions) are indicated.

As noted in Asempapa, R. et al [1], for COVID-19, values of certain parameters in the model, such as $k_p, k_e, \sigma_l, \rho_p$ can be obtained from related epidemiological studies, while others, including $\beta_p, \beta_h, \epsilon_p, \epsilon_h$, can be estimated by fitting the model to available data. Specifically, the default parameter values referenced from previous studies are summarized in Table 2 (units for all parameters in Table 2 are day-1), whereas the fitted parameter values, based on data from the first 120 days of the COVID-19 pandemic data in Brazil ($N = 212,615,886, I_l(0) = 1, E_l(0) = 2100, S_h(0) = 30\%, I_h(0) = 0, H(0) = 0$, and $R(0) = 0$) are presented in Table 3.

Parameter	Description
β_l	Effective community contact rate of low-risk susceptible individuals
β_h	Effective community contact rate of high-risk susceptible individuals
ν	Scaling factor for hospitalized transmission
ϵ_l	Non-pharmaceutical interventions by low-risk to reduce transmission
ϵ_h	Non-pharmaceutical interventions by high-risk to reduce transmission
R_0	Basic reproduction number
$k_l = k_h$	Rate of progression from the exposed class to the infected class
$l/k_l = l/k_h$	Average time from exposure to infection (latent period)
ρ_l	Proportion of infected that progress to hospital in the low-risk group
ρ_h	Proportion of infected that progress to hospital in the high-risk group
σ_l	Rate of progression from infected low-risk class to either hospitalized or recovery class
σ_h	Rate of progression from infected high-risk class to either hospitalized or recovery class
$\sigma_l \rho_l$	Rate of progression from infected low-risk class to hospitalized class
$\sigma_h \rho_h$	Rate of progression from infected high-risk class to hospitalized class
$\sigma_l (1-\rho_l)$	Rate of progression from infected low-risk class to recovery class
$\sigma_h (1-\rho_h)$	Rate of progression from infected high-risk class to recovery class
α	Recovery rate for hospitalized individuals
δ_l	Disease induced death rate of low-risk infected individuals
δ_h	Disease induced death rate of high-risk infected individuals
δ	Disease induced death rate of hospitalized individuals

Table 2. Default parameter values used in the deterministic and stochastic models (day⁻¹, unless noted). Parameter values adopted from prior epidemiologic studies and used as fixed inputs for model evaluation [5–7]. These values provide the baseline parameterization for simulations and for fitting procedures used to obtain Table 3.

Parameter	$k_l = k_h$	$\sigma_l \rho_l$	$\sigma_h \rho_h$	ρ_h	α	δ	δ_l	δ_h
Default Values	1/4	1/7	1/7	0.25	1/14	0.046	0.01	0.02

Table 3. Fitted parameter values for Brazil COVID-19 incidence data (first 120 days) and resulting R_0 . Estimated parameter values obtained by fitting the deterministic model to publicly available Brazil incidence data (Johns Hopkins University) over the first 120 days of the epidemic. The implied basic reproduction number R_0 is computed using Eq. (2) under the fitted baseline intervention levels ϵ_l and ϵ_h .

Parameter	β_l	β_h	ϵ_l	ϵ_h	ν	R_l	R_h	R_0
Default Values	0.9306	0.3361	0.2729	0.3432	0.1316	1.2554	0.1463	1.4017

Stochastic Continuous Time Markov Chain

We now introduce stochasticity into the model to capture stochastic variability in epidemic spread. The first step in this process is to recast the deterministic framework as a continuous-time Markov chain (CTMC) [8, 9]. This requires enumerating every possible state transition that can occur during an infinitesimal time interval and specifying the corresponding transition probability. Drawing directly from the compartmental structure, the resulting transition types and their associated rates are summarized in Table 4. These

rates constitute the infinitesimal generator of the CTMC, which we subsequently use to derive branching-process approximations.

Table 4. CTMC transition types and intensities (infinitesimal generator) for the stochastic epidemic model. All allowable state transitions of the CTMC correspond to the compartmental structure in Figure 1, together with their transition intensities. These intensities define the CTMC generator used for Doob–Gillespie simulation and for deriving the multi-type branching process approximation.

Description	State Transition	Rate
Low-risk Exposed individual	$S_l, E_l \rightarrow S_l - 1, E_l + 1$	$\beta_l(1 - \epsilon_l)(I_l(t) + I_h(t) + \nu H(t)) \frac{S_l(t)}{N}$
Low-risk Infected individual	$E_l, I_l \rightarrow E_l - 1, I_l + 1$	$k_l E_l(t)$
How-risk Exposed individual	$S_h, E_h \rightarrow S_h - 1, E_h + 1$	$\beta_h(1 - \epsilon_h)(I_l(t) + I_h(t) + \nu H(t)) \frac{S_h(t)}{N}$
High-risk Infected individual	$E_h, I_h \rightarrow E_h - 1, I_h + 1$	$k_h E_h(t)$
Low-Risk Infected individual Hospitalized	$I_l, H \rightarrow I_l - 1, H + 1$	$\sigma_l \rho_l I_l(t)$
High-risk Infected individual Hospitalized	$I_h, H \rightarrow I_h - 1, H + 1$	$\sigma_h \rho_h I_h(t)$
Low-risk Infected Recovery	$I_l, R \rightarrow I_l - 1, R + 1$	$\sigma_l(1 - \rho_l) I_l(t)$
High-risk Infected Recovery	$I_h, R \rightarrow I_h - 1, R + 1$	$\sigma_h(1 - \rho_h) I_h(t)$
Hospitalized Recovery	$H, R \rightarrow H - 1, R + 1$	$\alpha H(t)$
Low-risk Death	$I_l, D \rightarrow I_l - 1, D + 1$	$\delta_l I_l(t)$
High-risk Death	$I_h, D \rightarrow I_h - 1, D + 1$	$\delta_h I_h(t)$
Hospitalized Death	$H, D \rightarrow H - 1, D + 1$	$\delta H(t)$

Using the Doob–Gillespie algorithm [10], we generate stochastic realizations of the infected compartments for both high-risk and low-risk subpopulations. The CTMC is calibrated to the Brazilian COVID-19 data with parameter values previously cited. While stochastic fluctuations appear negligible with such a big population size, the variability becomes evident when zooming up the trajectories,

or when the model is run for a smaller populations size. To better visualize the differences between the infection trajectories predicted by the deterministic model and those obtained from the stochastic realizations of the CTMC model for both subpopulations, Figure 2 is generated using a smaller population size ($N=2,000$).

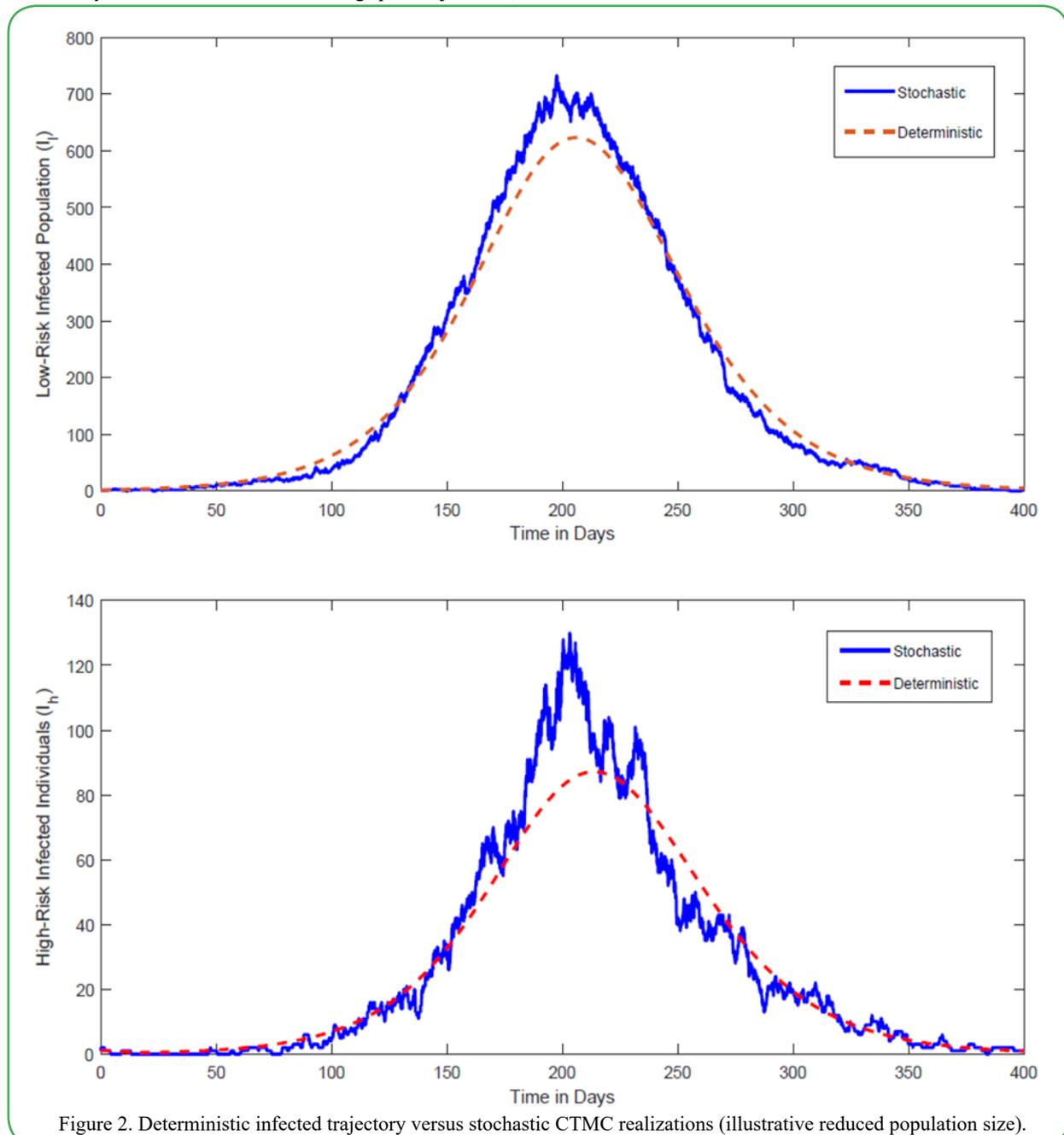


Figure 2. Deterministic infected trajectory versus stochastic CTMC realizations (illustrative reduced population size).

Comparison of the deterministic infected trajectory from system (1) with representative stochastic realizations from the CTMC (simulated via the Doob–Gillespie algorithm) under the same parameterization. For visualization, a reduced population size N is used to make stochastic variability and stochastic extinction events apparent. Curves show stochastic paths fluctuating around the deterministic trajectory; extinction corresponds to paths reaching zero infected individuals.

As illustrated in Figure 2, the stochastic realizations fluctuate above or below the deterministic trajectory, reflecting the inherent randomness in epidemic dynamics in real-world settings. In particular, a fraction of the stochastic paths dies out completely, even when the basic reproduction number $R_0 > 1$; this phenomenon, known as *stochastic extinction*, is a feature that deterministic models cannot capture.

As shown by the deterministic model, when $R_0 > 1$, deterministic models predict an inevitable epidemic, although earlier interventions can slow the spread, reduce but broaden the peak, and postpone its arrival. In contrast, real-world evidence shows that appropriately sized public-health prevention measures can curb the epidemic even when $R_0 > 1$. Below, we run stochastic simulations to estimate the stochastic extinction probability for $R_0 > 1$ with various combinations of early intervention strength levels in the low-risk and high-risk populations.

Starting from one infectious individual in the high-risk population and none in the low-risk population, the simulated disease extinction probabilities as a function of the health policy intervention parameters for the low-risk group (ϵ_l) and that for the high-risk group (ϵ_h), by averaging 10,000 realizations of the CTMC, are shown in Table 5.

In Table 5 and the subsequent Table 6 and Table 7, same as in Figure 4, the values shown in the ϵ_l and ϵ_h columns represent multiplicative scaling factors relative to their respective baseline values, rather than the actual parameter values. The baseline values used are $\epsilon_l = 0.2729$ and $\epsilon_h = 0.3432$, as specified in Section 2.1. Therefore, for ϵ_l , the baseline value of 0.2729 corresponds to a scaling factor of 1, while scaling factors of 1.4, 1.6, and 1.8 correspond to actual ϵ_l values of 0.3821, 0.4366, and 0.4912, respectively. Similarly, for ϵ_h , the baseline value of 0.3432 corresponds to a scaling factor of 1, and scaling factors of 1.4, 1.6, and 1.8 correspond to actual ϵ_h values of 0.4805, 0.5491, and 0.6178, respectively. It can be verified easily that for all combinations of ϵ_l and ϵ_h values appeared in Tables 5-7, the corresponding basic reproduction number R_0 is greater than 1.

$I_l(0)$	$I_h(0)$	ϵ_l	ϵ_h	Estimated Extinction Probability
0	1	-	-	67.2190%
0	1	1.4	1	74.9339%
0	1	1.6	1	77.0346%
0	1	1.8	1	81.0497%
$I_l(0)$	$I_h(0)$	ϵ_l	ϵ_h	Estimated Extinction Probability
0	1	-	-	67.2190%
0	1	1	1.4	68.4338%
0	1	1	1.6	68.9118%
0	1	1	1.8	69.5878%
$I_l(0)$	$I_h(0)$	ϵ_l	ϵ_h	Estimated Extinction Probability
0	1	-	-	67.2190%
0	1	1.4	1.4	75.1317%
0	1	1.6	1.6	77.2753%
0	1	1.8	1.8	82.1869%

From Table 5, when the epidemic is seeded by a single infectious individual in the high-risk group (and no infection in the low-risk group), the stochastic extinction probability under baseline intervention levels ($\epsilon_l = 0.2729$, $\epsilon_h = 0.3432$, $R_0 = 1.4017$), is roughly 67%. Increasing either ϵ_l , or ϵ_h , or both raises the extinction probability, but the magnitude of the effect differs substantially. Specifically, increasing ϵ_h alone leads to a slight improvement in the extinction probability, whereas an equivalent proportional increase in ϵ_l produces a much larger rise in the extinction probability, nearly matching the impact of increasing both ϵ_l and ϵ_h simultaneously at the same scale.

For example, when ϵ_l is held at its baseline and ϵ_h is increased to $1.4 \times \text{baseline}$ (a 40% increase from its baseline value), the extinction probability increases only from roughly 67% to about 68%. In contrast, when $\epsilon_l = 1.4 \times \text{baseline}$ (with ϵ_h at its baseline value), or when both ϵ_l and ϵ_h are increased to $1.4 \times \text{baseline}$, the extinction probability rises to roughly 75%.

These results indicate that even when the epidemic originated solely in the high-risk subpopulation, strengthening interventions in the low-risk groups can yield an extinction probability comparable to that achieved by intensifying interventions of the same scale to both groups. By contrast, focusing exclusively on the high-risk group produces a substantially lower likelihood of disease extinction.

In the following section, we discuss how to estimate the probability of stochastic extinction using a branching process approximation.

Multi-type branching process approximation

An analytical approach of estimating the probability of stochastic extinction can be obtained from branching process theory. The idea of branching processes dates back to the French actuary Irenée-Jules Bienaymé, whose 1845 paper “*On the Law of Multiplication and the Duration of Families*” [11] anticipated the well-known

Table 5. Estimated extinction probabilities from CTMC simulation under intervention-scaling scenarios (initial seeding in high-risk group).

Estimated epidemic extinction probabilities computed by averaging 10,000 Doob–Gillespie realizations of the CTMC under alternative intervention strengths. Initial condition: one infectious individual in the high-risk group and none in the low-risk group ($I_h(0) = 1$, $I_l(0) = 0$; other infectious compartments initially zero unless otherwise stated). Columns labeled ϵ_l and ϵ_h report multiplicative scaling factors applied to the baseline fitted values in Table 3 (e.g., 1.0 = baseline; 1.4 = 40% increase). For each scenario, $R_0 > 1$.

Bienaymé-Galton-Watson (BGW) branching process. The original BGW model was devised to study the extinction of family names under the assumption that surnames are transmitted solely from father to son, and all men have the same probability of having 0, 1, 2, ... sons reaching adulthood. Thus, a family name becomes extinct when all male descendants die without adult male heirs. After a period of dormancy, research on branching processes revived in the early 20th century and expanded substantially after World War II. In this study, we employ a multi-type branching process to approximate our stochastic epidemic model.

Specifically, in our CTMC model, five infectious compartments are considered:

- Class 1: low-risk exposed individuals (E_l)
- Class 2: low-risk infectious individuals (I_l)
- Class 3: high-risk exposed individuals (E_h)
- Class 4: high-risk infected individuals (I_h)
- Class 5: Hospitalized individuals (H)

Accordingly, in the multi-type branching process framework, the process is represented as

$$\mathbf{V}(t) = (E_l(t), I_l(t), E_h(t), I_h(t), H(t)),$$

where $E_l(t)$, $I_l(t)$, $E_h(t)$, $I_h(t)$ and $H(t)$ denote the numbers of individuals in the corresponding infectious class at time t . The offspring distribution can be derived directly from the state transition rates provided in Table 4. Because each transition is Poisson in continuous time, the numbers of offspring of the five types are independent Poisson variables with means equal to the corresponding rates divided by the total exit rate of the parent.

For each type i ($i = 1, 2, \dots, 5$), we define the probability-generating function (pgf),

$$G_i(\mathbf{s}) = E[s_1^{X_{i1}} s_2^{X_{i2}} \dots s_5^{X_{i5}}], \mathbf{s} = (s_1, \dots, s_5),$$

where X_{ij} is the number of type- j offspring produced by a single type- i . The vector $\boldsymbol{\pi} = (\pi_1, \dots, \pi_5)$ satisfies the fixed-point system

$$\pi_i = G_i(\pi_1, \pi_2, \pi_3, \pi_4, \pi_5), i = 1, \dots, 5$$

The overall extinction probability for an arbitrary initial state $\mathbf{k} = (k_1, \dots, k_5)$ is then

$$\Pi(\mathbf{k}) = \prod_{i=1}^5 \pi_i^{k_i},$$

i.e. the product of the single-type extinction probabilities raised to the number of initial individuals of each type.

Hence, let $E_i(0)$, $I_i(0)$, $E_h(0)$, $I_h(0)$, $H(0)$ denote the initial counts of individuals across the five infectious classes, and $\mathbf{0} = (0, 0, 0, 0, 0)$, the overall extinction probability is

$$P_{\text{extinction}} = P\{\mathbf{V}(t) = \mathbf{0}\} = \pi_1^{E_i(0)} \cdot \pi_2^{I_i(0)} \cdot \pi_3^{E_h(0)} \cdot \pi_4^{I_h(0)} \cdot \pi_5^{H(0)} \quad (3)$$

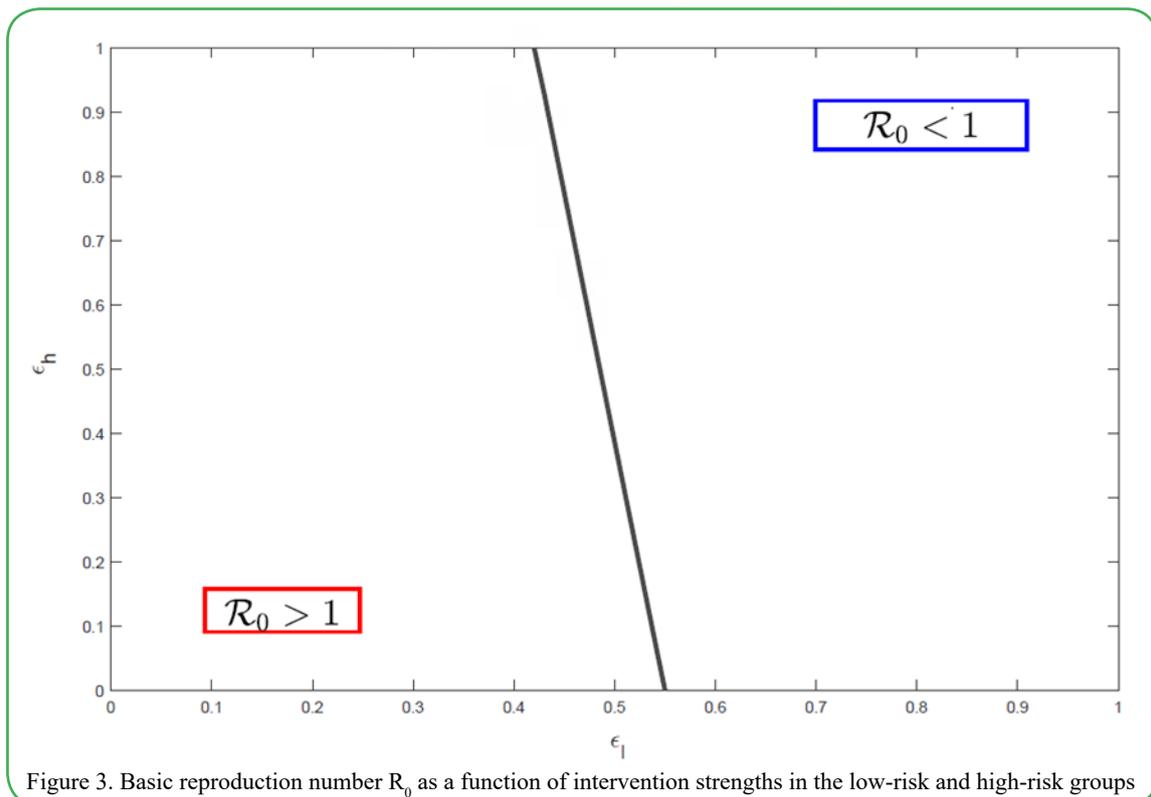


Figure 3. Basic reproduction number R_0 as a function of intervention strengths in the low-risk and high-risk groups

Contour plot of R_0 computed from Eq. (2) across values of the intervention parameters ϵ_i (low-risk) and ϵ_h (high-risk), holding all other parameters fixed at baseline/fitted values (Tables 2–3). The threshold $R_0 = 1$ is indicated to delineate deterministic growth versus decline.

Though equation (2) shows that, holding all other parameters constant, R_0 is a linear combination of ϵ_i and ϵ_h , and Figure 3 illustrates how these two interventions rates influence R_0 . The results highlight that the non-pharmaceutical intervention level in the low-risk group, ϵ_i , is the decisive factor for disease extinction. Specifically:

- When $\epsilon_i < 0.41$: no degree of interventions in the high-risk group can reduce R_0 below 1, implying that the disease will persist regardless of efforts targeted solely at the high-risk population.
- When $\epsilon_i > 0.55$: R_0 can fall below 1 even in the absence of interventions in the high-risk group, ensuring that the disease will eventually die out.

Solutions to the deterministic model (1) also depend on ϵ_i and ϵ_h . Holding all other parameters at their values listed in Tables 2 and 3, and specifically, ϵ_h is fixed at its baseline value of 0.3432, Figure 4 illustrates how the Susceptible and Exposed populations evolve in both the low-risk and high-risk populations as ϵ_i varies. Here, the

Results

Deterministic baseline dynamics

Asempapa, R. et al [1] demonstrated that the cumulative COVID-19 cases predicted by this compartmental model for the first 120 day of the epidemic in Brazil match the reported data closely, with only minor random deviations. This concordance validates the model as a robust baseline for embedding stochastic components that capture random fluctuations.

As we can see in Table 3, for the COVID-19 outbreak in Brazil, the estimated value of the basic reproduction number is $R_0 = 1.4$. This means that, on average, each infectious individual transmits the disease to roughly 1.4 others, so the epidemic is expected to grow, making an outbreak unavoidable under the baseline non-pharmaceutical intervention levels in the low-risk and high-risk groups $\epsilon_i = 0.2729$ and $\epsilon_h = 0.3432$, respectively.

baseline value of ϵ_i is 0.2729 (as listed in Table 3), and numbers in the legend denote multiplicative factors relative to this baseline (for example, “(0.5) baseline” means $\epsilon_i = 0.5 \times 0.2729 = 0.1365$). Thus, multipliers less than 1 represent relaxed health prevention measures, whereas multipliers greater than 1 indicate stricter health intervention policies compared to the default baseline intervention conditions.

Deterministic solutions of system (1) showing the evolution of susceptible and exposed populations in the low-risk and high-risk subpopulations as ϵ_i varies by multiplicative factors relative to its baseline fitted value (Table 3), while holding ϵ_h fixed at its baseline fitted value. Legend entries denote the multiplier applied to baseline ϵ_i (e.g., 1.0 = baseline). All other parameters are fixed at Tables 2–3.

The trajectories in Figure 4 reveal that stronger interventions in the low-risk population (higher ϵ_i) slow the decline of the susceptible population, reduce the magnitude of the exposed (and thus infected) peak, and delay the peak’s occurrence in both the low-risk and high-risk subpopulation. Conversely, weaker interventions (smaller ϵ_i) accelerate the depletion of susceptible individuals and yield a higher and earlier exposed (and thus infected) peak in both groups. When ϵ_i is increased to 1.8 times its baseline value, the susceptible population in both the low-risk and high-risk groups remains nearly constant, and the disease exerts minimal impact on either subpopulation.

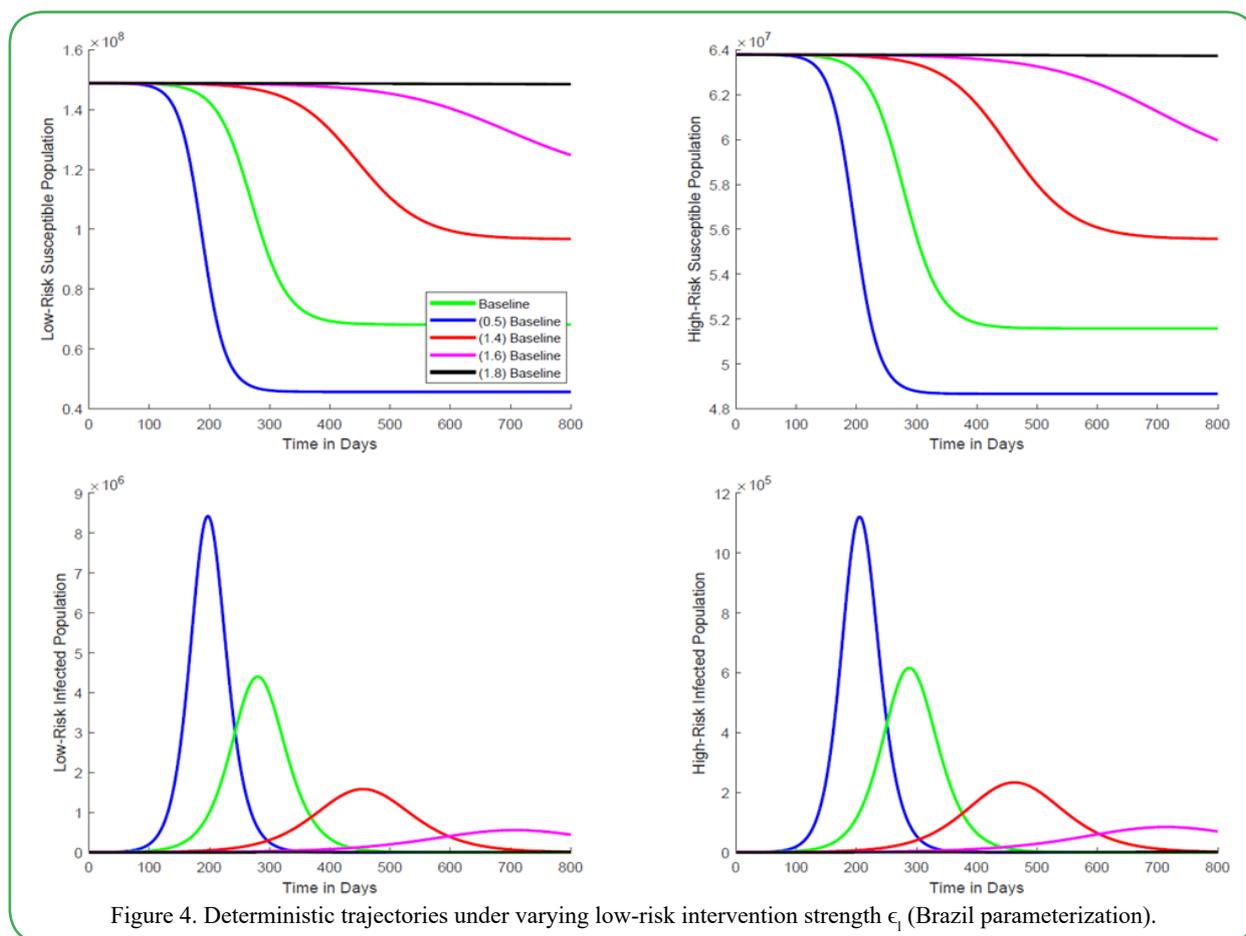


Figure 4. Deterministic trajectories under varying low-risk intervention strength ϵ_l (Brazil parameterization).

Stochastic extinction probabilities and scenario comparisons

In this section, we evaluate the stochastic extinction probabilities for the two different initial scenarios mentioned in the Methods section. The purpose is twofold: to validate the branching-process calculations against the numerical simulation of the CTMC, and to explore how extinction probabilities change under alternative initial conditions and different levels of interventions.

As mentioned previously, all parameter combinations listed in Tables 6–7 satisfy $R_0 > 1$, thus the deterministic models predict an inevitable outbreak. Nevertheless, the stochastic framework reveals non-zero extinction probabilities.

First, we consider the case discussed in the baseline initial condition: a single initially infected individual occurs in the high-risk population and no infections in the low-risk population. In the baseline initial condition, the stochastic extinction probabilities were estimated by averaging 10,000 Gillespie realizations of the CTMC model for each combination of ϵ_l and ϵ_h values. Here, we compute the extinction probabilities directly from the fixed-point system (3). These results are summarized in Table 6.

From Table 6, it is evident that the extinction probabilities obtained using two approaches are very close, confirming the reliability of both. These findings further support the conclusion presented in the baseline initial condition: even when the single initial infection occurs in the high-risk group and the low-risk group has no infections, implementing stricter non-pharmaceutical interventions in the low-risk group alone is nearly as effective as applying them to both groups, and substantially more effective than targeting only the high-risk group.

Table 6. Validation of extinction probabilities: CTMC simulation versus branching-process calculation (initial seeding in high-risk group). Comparison of extinction probabilities estimated by CTMC simulation (10,000 Gillespie realizations) versus extinction probabilities computed from the multi-type branching process fixed-point equations under the same intervention-scaling scenarios. Initial condition corresponds to Table 5 ($I_h(0)=1, I_l(0)=0$). Scaling factors in ϵ_l and ϵ_h are relative to baseline fitted values in Table 3.

Next, we examine an alternative initial situation: a single initially infected individual occurs in the low-risk population with no infections in the high-risk population. Again, we estimate the disease extinction probability by both the simulation approach and the branching process approach for different combinations of the ϵ_l and ϵ_h values. The results are presented in Table 7.

Table 7 shows similar patterns to those observed in the first scenario regarding how the level of non-pharmaceutical interventions in the low-risk and high-risk groups (ϵ_l and ϵ_h , respectively) influence the stochastic disease extinction probabilities. Specifically, under baseline intervention levels for both groups ($\epsilon_l = 0.2729, \epsilon_h = 0.3432, R_0 = 1.4017$), the estimated stochastic disease extinction probability for this scenario is approximately 68%. As the strength of non-pharmaceutical interventions increases in either group or both groups, the extinction probability rises accordingly. However, the magnitude of the impact on the extinction probability differs between ϵ_l and ϵ_h . Once again, applying stricter interventions in the low-risk population alone achieves nearly the same extinction probability as applying them to both populations, and performs markedly better than focusing interventions exclusively on the high-risk population.

I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
0	1	-	-	67.2190%	67.1837%
0	1	1.4	1	74.9339%	74.1167%
0	1	1.6	1	77.0346%	77.6754%
0	1	1.8	1	81.0497%	81.6792%
I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
0	1	-	-	67.2190%	67.1837%
0	1	1	1.4	68.4338%	68.3423%
0	1	1	1.6	68.9118%	69.1589%
0	1	1	1.8	69.5878%	70.0114%
I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
0	1	-	-	67.2190%	67.1837%
0	1	1.4	1.4	75.1317%	76.4470%
0	1	1.6	1.6	77.2753%	77.9367%
0	1	1.8	1.8	82.1869%	82.9794%

Table 7. Extinction probabilities under alternative initial seeding (initial infection in low-risk group): simulation versus branching-process calculation. Extinction probabilities estimated by CTMC simulation (10,000 Gillespie realizations) and computed using the branching-process approximation under alternative intervention-

scaling scenarios. Initial condition: one infectious individual in the low-risk group and none in the high-risk group ($I_l(0)=1, I_h(0)=0$; other infectious compartments initially zero unless otherwise stated). Scaling factors in ϵ_l and ϵ_h are relative to baseline fitted values in Table 3; for each scenario, $R_0 > 1$.

I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
1	0	-	-	67.7710%	67.7783%
1	0	1.4	1	73.1319%	73.1401%
1	0	1.6	1	76.8723%	76.8711%
1	0	1.8	1	80.5894%	80.5923%
I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
1	0	-	-	67.7710%	67.7783%
1	0	1	1.4	68.7642%	69.4532%
1	0	1	1.6	69.2665%	69.7505%
1	0	1	1.8	69.9423%	69.9948%
I_l	I_h	ϵ_l	ϵ_h	Estimated Extinction Probabilities	Calculated Extinction Probabilities
1	0	-	-	67.7710%	67.7783%
1	0	1.4	1.4	75.0713%	75.7634%
1	0	1.6	1.6	77.7876%	77.2918%
1	0	1.8	1.8	82.5660%	82.8120%

Discussion

Extinction probability provides a decision-relevant complement to deterministic metrics such as the basic reproduction number. In early epidemic phases, when case counts are low, random variation in contacts and transmission can produce a non-trivial chance of fade-out even when a deterministic model with fitted parameters implies growth ($R_0 > 1$). By embedding a continuous-time Markov chain (CTMC) within a standard compartmental framework and approximating early dynamics using a multi-type branching process, we quantify how alternative intervention strengths shift the chance

that transmission chains terminate before sustained epidemic growth occurs.

These results fit within the broader public health modeling literature. Deterministic compartmental models remain essential for describing average trajectories and for interpreting how parameters shape dynamics [2-4]. However, stochastic epidemic models are required to characterize extinction versus major outbreak outcomes when infections are sparse, and CTMC formulations and branching-process approximations provide standard tools for this purpose [5, 6]. In the COVID-19 context, early estimates of transmission potential

and analyses of non-pharmaceutical interventions underscore that growth conditions alone do not determine realized outcomes and that timely suppression of community transmission can materially change epidemic trajectories [8-10].

A central finding is that proportional strengthening of non-pharmaceutical interventions in the low-risk subpopulation produces substantially larger gains in extinction probability than an equivalent proportional increase confined to the high-risk subpopulation. This pattern is consistent with the principle that suppressing transmission in the larger transmitting group reduces the force of infection experienced by all subpopulations, including high-risk individuals, thereby increasing the likelihood of early fade-out even when the outbreak is initially seeded in the high-risk group.

From a public health implementation standpoint, the results emphasize broad measures that reduce effective contacts and improve adherence in settings that drive transmission. For respiratory epidemics, actionable levers include reducing crowding in high-contact venues, improving indoor air quality, promoting masking during periods of elevated risk, enabling rapid access to testing and isolation, and supporting adherence through clear risk communication and practical resources that make compliance feasible. Targeted protections for high-risk groups remain essential for reducing severe outcomes, but our results indicate that high-risk-only strategies are unlikely to maximize the probability of extinction when substantial transmission potential persists in the broader community.

Finally, the close agreement between extinction probabilities estimated via Monte Carlo CTMC simulation and those computed from the branching-process fixed-point system supports the internal validity of the analytic approximation under the early-epidemic conditions considered. Practically, extinction probability can be used as an explicit endpoint to compare candidate intervention packages, to justify early population-level measures when elimination remains feasible, and to communicate probabilistic benefits to decision-makers when deterministic predictions alone suggest inevitable growth.

Limitations

This work has several limitations that should be considered when interpreting the findings and translating them to policy.

- The branching-process approximation is most accurate during the early outbreak phase; accuracy may diminish as susceptible depletion becomes substantial or as behavioral responses and time-varying interventions alter transmission dynamics over time.
- The model assumes homogeneous mixing within and between the low-risk and high-risk groups. Real-world contact patterns are heterogeneous and structured by age, household composition, occupation, geography, and network effects, which could change the relative impact of targeted versus population-wide interventions.
- Interventions are represented through constant multiplicative reductions in effective transmission. In practice, NPIs are implemented and relaxed over time, adherence varies across individuals and settings, and enforcement and risk perception change dynamically.
- Parameter uncertainty is not fully propagated. Parameters are estimated from early-epidemic incidence data, and different parameter combinations can produce similar fitted trajectories, which can affect derived extinction probabilities and comparisons between scenarios.
- The analysis uses publicly available, aggregated incidence data. Such data can be affected by reporting delays, under-ascertainment, changes in testing capacity and case definitions, and temporal aggregation, particularly during the first months of an epidemic.

- The model does not explicitly incorporate vaccination, variants, therapeutics, imported cases, or spatial spread. Therefore, results are most directly applicable to early-stage outbreaks in largely susceptible populations under relatively stable transmission conditions.

Public Health Significance and Policy Implications and Recommendations

These results support several actionable public health implications for early outbreak response, particularly when elimination remains feasible and surveillance indicates low-to-moderate incidence.

First, intervention packages intended to increase the probability of extinction should prioritize measures that reduce transmission broadly in the community, especially those that improve compliance or reduce contacts in the larger low-risk group. Examples include interventions in high-contact environments (workplaces, schools, and mass gatherings), clear risk communication and adherence support, and rapid testing and isolation infrastructure that enables early interruption of transmission chains.

Second, targeted protections for high-risk groups should be implemented in parallel to reduce morbidity and mortality, but they should be viewed as insufficient on their own for maximizing extinction probability when substantial transmission potential persists in the broader community.

Third, extinction probability can be operationalized as a decision metric alongside R_0 and incidence trends. Agencies can use the framework to compare candidate non-pharmaceutical intervention packages by quantifying how much each package increases the chance of elimination over a defined time horizon, thereby enabling transparent comparison of intervention intensity versus expected benefit.

Finally, these findings underscore the value of timeliness. When incidence is low, short, decisive population-wide measures may yield disproportionately large gains in extinction probability relative to later implementation, when transmission chains are numerous and stochastic fade-out is less likely.

- Use extinction probability to rank candidate non-pharmaceutical intervention packages during the early response window, rather than relying on R_0 alone.
- Prioritize population-wide transmission suppression measures that are feasible to implement rapidly and that sustain high adherence in the broader low-risk population.
- Maintain targeted protections for high-risk groups to reduce severe outcomes, but avoid relying exclusively on high-risk-only interventions to achieve elimination.

Conclusions

Although when the basic reproduction number exceeds 1, the deterministic model alone would forecast an inevitable epidemic, our incorporation of stochasticity through a CTMC and a multi-type branching process approximation, demonstrates that non-pharmaceutical interventions can still produce a non-negligible probability of disease extinction, in line with empirical observations.

The analysis pinpoints compliance in the low-risk subpopulation as the key lever for maximizing extinction probability. Strategies that focus solely on the high-risk group are insufficient; in contrast, elevating intervention levels in the low-risk subpopulation yields substantially greater benefits. While strengthening interventions in both subpopulations provides the highest extinction probability, the improvement is nearly comparable to targeting the low-risk group alone.

These findings underscore the importance of stochastic modeling for public-health decision-making and suggest that, in future outbreaks, resources should be directed preferentially toward broad, population-

wide interventions, especially those aimed at the low-risk group, to maximize the likelihood of epidemic control.

Conflict of Interest Statement: The authors have no conflicts of interest to disclose.

Author Contributions

All authors contributed equally to all aspects of this manuscript. All authors approved the final manuscript and agree to be accountable for all aspects of the work.

Ethics Approval/Consent

IRB review and informed consent were not required because the study used publicly available, de-identified, aggregated incidence data.

Funding Statement

The authors received no internal or external funding for this work.

Data Availability Statement

All data used in this study are publicly available, de-identified, aggregated incidence data.

References

1. Asempapa, R., Oduro, B., Apenteng, O. O. & Magagula, V. M. A COVID-19 mathematical model of at-risk populations with non-pharmaceutical preventive measures: The case of Brazil and South Africa. *Infect. Dis. Model.* 7, 45–61 (2021).
2. Blackwood, J. C. & M.Childs, L. An introduction to compartmental modeling for the budding infectious disease modeler. *Lett. Biomath.* 5, 195–221 (2018).
3. Hamer, W. H. The Milroy Lectures on Epidemic Disease in England — The Evidence of Variability and of Persistency of Type. *The Lancet*, 167, 569–574 (1906).
4. Brauer, F. Mathematical epidemiology: Past, present, and future. *Infect. Dis. Model.* 2, 113–127 (2017).
5. Li, Q. et al. Early Transmission Dynamics in Wuhan, China, of Novel Coronavirus–Infected Pneumonia. *N. Engl. J. Med.* 382, 1199–1207 (2020).
6. Tang, B. et al. Estimation of the Transmission Risk of the 2019-nCoV and Its Implication for Public Health Interventions. *J. Clin. Med.* 9, 462 (2020).
7. Ferguson, N. et al. Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. <https://doi.org/10.25561/77482> (2020) doi:10.25561/77482.
8. Allen, L. J. S. A primer on stochastic epidemic models: Formulation, numerical simulation, and analysis. *Infect. Dis. Model.* 2, 128–142 (2017).
9. Lahodny Jr., G. E., Gautam, R. & Ivanek, R. Estimating the probability of an extinction or major outbreak for an environmentally transmitted infectious disease. *J. Biol. Dyn.* 9, 128–155 (2015).
10. Gillespie, D. T. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comput. Phys.* 22, 403–434 (1976).
11. Bienaymé, J. De la loi de multiplication et de la durée des familles. *Soc Philomat. Paris Extraits*, 13, 131-132.