

Article

A Genotype-Structured PDE Model of Vaccine-Induced Control of Variant Emergence

Anass Bouchnita and Behzad Djafari-Rouhani * 

Department of Mathematical Sciences, The University of Texas at El Paso, El Paso, TX 79968, USA;
abouchnita@utep.edu

* Correspondence: behzad@utep.edu

Abstract

Variant emergence continues to pose a threat to public health, despite the widespread use of vaccination. To quantify how vaccine strain compositions shape evolutionary and epidemiological outcomes, we extend a previous genotype-structured transmission model with vaccination and study the impact of different vaccination formulations on variant emergence. It consists of a set of partial differential equations coupled with an integro-differential one. We begin by showing that the model reproduces variant emergence followed by a period of co-circulation in the absence of vaccination. Then, we introduce vaccination and show important trade-offs shaped by the breadth and cross-protection of vaccine-induced immunity. In our simulations, narrow-spectrum vaccines substantially reduce the immediate infection burden but inadvertently promote the emergence of non-targeted variants. After that, we study the effects of more complex shapes such as triangular and M-shaped configurations. We show that M-triangular distributions outperform triangular ones by limiting secondary variant expansion for vaccines with narrow cross-protection. In contrast, triangular compositions are more protective when considering broader cross-protection. We also show that targeting the genetic area between co-circulating variants is more beneficial than focusing on specific variants when using vaccines with a broad cross-protection. Together, these results highlight how vaccine breadth and antigenic targeting influence both epidemic size and the trajectory of variant emergence, offering quantitative guidance for monovalent and multivalent vaccine design.

Keywords: genotype-structured modeling; infectious disease modeling; vaccine formulation; numerical simulations; partial differential equations

MSC: 37N25; 93C20; 45K05



Academic Editor: Ruizhi Yang

Received: 3 December 2025

Revised: 24 January 2026

Accepted: 31 January 2026

Published: 11 February 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Strain selection for monovalent and multivalent vaccines relies on a coordinated laboratory workflow that integrates global surveillance, antigenic characterization, and immunological testing. Viral isolates collected through systems such as the Global Influenza Surveillance and Response System (GISRS) or genomic surveillance programs undergo sequencing and antigenic profiling to map genetic and antigenic diversity across circulating lineages. Neutralization assays—including hemagglutination inhibition (HI) for influenza, microneutralization assays, pseudovirus neutralization tests, and monoclonal antibody escape profiling—quantify antigenic distances and identify strains representing distinct antigenic clusters [1–4]. These data produce “antigenic maps” that guide the selection

of candidate strains that span broad regions of antigenic space, minimizing gaps where immune escape variants could emerge.

Candidate strains are next evaluated through *in vivo* immunogenicity and challenge studies to determine cross-protection breadth, immunodominance patterns, and potential immunological interference within multivalent combinations. Animal models such as mice, ferrets, hamsters, or nonhuman primates are used to compare single-strain versus multi-strain immunization, assessing neutralizing titers, T-cell responses, viral load after challenge, and pathology [5,6]. Parallel experiments test manufacturability and genetic stability, ensuring that selected strains grow efficiently in production systems and maintain antigenic fidelity.

The rationale for selecting specific strains in cocktail vaccines is to maximize antigenic coverage while minimizing the evolutionary opportunities available for viral escape. Because rapidly evolving respiratory viruses such as influenza, SARS-CoV-2, and RSV accumulate mutations that alter antigenic profiles, combining strains from distinct antigenic or genetic clusters improves the breadth of neutralizing responses and reduces sensitivity to antigenic drift [2,7]. By targeting antigenically distant variants simultaneously, multivalent formulations exploit cross-protective immunity and create immunological “barriers” that constrain viable mutational pathways for the virus [8]. Empirical studies have shown that broad-spectrum immunity generated by multivalent vaccines slows adaptation and lowers the probability of escape mutants emerging during transmission, providing a robust evolutionary rationale for cocktail design [9,10]. This strategy is especially important for pathogens with high mutation rates and complex genotype–phenotype landscapes, where single-strain vaccines risk narrowing immune pressure and inadvertently promoting variant emergence.

Various mathematical modeling frameworks have been employed to guide vaccine formulation and evaluation. Within-host viral–immune models have been used to characterize antibody development after vaccination, incorporating mechanisms such as B-cell differentiation [11], T-cell-mediated responses [12], and dose–response effects [13]. At the population level, transmission models have been extended to assess the impact of vaccination on epidemic dynamics and to support public-health decision-making [14–16]. Such models have been optimized to derive age-specific vaccination strategies during the COVID-19 pandemic [17] and fitted to surveillance data to estimate vaccine effectiveness [18]. Multi-strain models have further elucidated how vaccination alters competition among co-circulating strains [19]. Integrating viral evolution with transmission and vaccination, other work has shown that conventional strain-matched vaccines can increase the likelihood of variant emergence, whereas broader universal vaccines suppress antigenic evolution [20].

To the best of our knowledge, no transmission-dynamics framework explicitly incorporates vaccine formulation and its impact on variant emergence. Here, we address this gap by extending a previously developed genotype-structured model with immunity to include vaccination effects [21,22]. From a mathematical perspective, our framework belongs to the class of structured population models, where heterogeneity is represented through continuous trait variables. Such models have been widely used to study evolutionary dynamics, selection, and adaptation using transport–diffusion equations and nonlocal interaction terms [23,24]. In epidemiology, genotype- or phenotype-structured models provide a natural bridge between population-level transmission dynamics and evolutionary processes [21,25].

In this framework, vaccines are represented as genotype-dependent functions, enabling the evaluation of narrow and broad formulations, including Gaussian, rectangular, triangular, and M-shaped profiles. We first simulate the baseline scenario in the absence of vaccination. We then introduce vaccine-mediated immunity and use numerical simulations

to quantify the number of averted infections as a function of vaccine breadth, formulation shape, cross-immunity properties, and targeted genotype location. Finally, we discuss the biological interpretation of these results and outline future research directions.

2. A Genotype-Structured Model of Disease Transmission with Vaccination

2.1. Model Description

We extend a previously developed genotype-structured transmission model to account for the effects of vaccination. As in our earlier models for COVID-19, influenza, and RSV [26,27], the framework introduces a state variable describing population-level immunity. This work builds directly on our past studies of variant emergence in the absence of vaccination [21,22], but here, we explicitly incorporate, and analyze, the impact of vaccine-induced immunity. Vaccination is introduced into the model as a genotype-dependent function, enabling the representation of customizable strain compositions and strain-targeted antigen profiles.

The model assumes that recovered individuals contribute to the overall population-immunity, which in turn modulates transmission and mortality rates. This formulation provides a flexible way to integrate immunological processes within compartment-based epidemiological models. The viral genotype, denoted by x , evolves in a one-dimensional genotype space X defined over the interval $[0, 10]$ in arbitrary genetic units (G.U.). These units serve as a metric to quantify genotype variation arising from host-level mutations and to characterize relationships between inter-variant genetic distance and cross-immunity. The model tracks susceptible (S), infected (I), and recovered (R) populations. Susceptibility represents a fully naïve immune state and is therefore assumed to be genotype-independent. In contrast, population-level immunity retains memory of past exposures and is structured by the genotype of the infecting or vaccine-targeted strains, reflecting the antigen-specific nature of immune memory. Mutational changes during replication are represented through a diffusion term in the infected equation, following the approach of [28]. Population-immunity is updated through both infection- and vaccine-derived components, described by genotype-stratified variables M and M_v . Infection-derived immunity increases with new recoveries, whereas vaccine-derived immunity grows with administered doses; both wane over time. Each immunity source confers genotype-specific cross-immunity, such that its effectiveness depends on the genetic distance between circulating strains and existing antibodies. Transmission is reduced according to this immunity and its corresponding efficacy. The overall model structure and the interactions between epidemiological compartments and immunity variables are summarized in Figure 1.

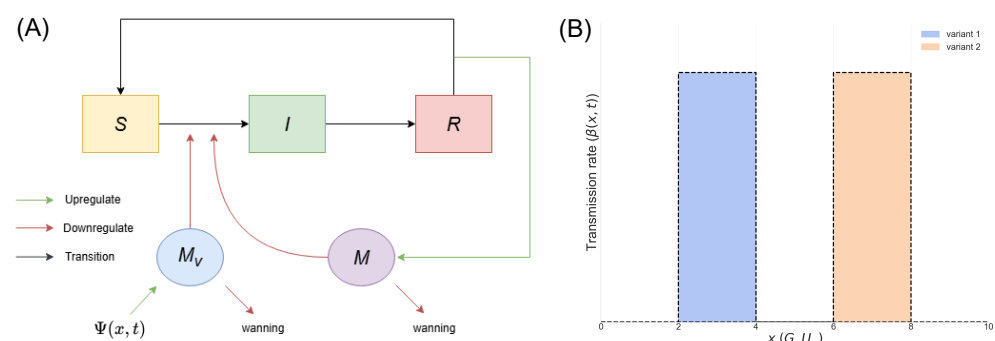


Figure 1. (A) A scheme showing the interactions between the compartments and the immunity variables involved in the model. (B) A representation of the fitness landscape for the transmission rate that is considered in the simulations to study variant emergence. It shows two viable spaces corresponding to two variants 1 and 2.

To study the competition and alternation between variants 1 and 2, we assume that the transmission rate follows a fitness landscape containing two influenza variants: influenza 1 and influenza 2 (Figure 1B). Each variant is represented as occupying a region of width 2 G.U., so that variant 1 spans the interval $[2, 4]$, where 3 denotes the center of the influenza 1 cluster. An analogous definition applies to variant 2. Both variants share the same intrinsic transmission rate, corresponding to the baseline reproduction number for seasonal influenza. The non-survival zone corresponds to genotypes with low fitness caused by constraints such as impaired protein folding, reduced receptor binding, inefficient replication, or strong cross-reactive immune neutralization. The overall framework follows standard SIRS dynamics, where the susceptible population is assumed to be independent of genotype and evolves according to:

$$\begin{aligned} \frac{dS}{dt}(t) = & -S(t) \int_X \beta(x) \frac{I(x, t)}{1 + K_1 \int_X \phi_x(y, d_0^1) M(y, t) dy + K_2 \int_X \phi_x(y, d_0^2) M_v(y, t) dy} dx \\ & + \delta \int_X R(x, t) dx. \end{aligned} \quad (1)$$

The first term on the right-hand side represents new infections among susceptible individuals. Here, $\beta(x, t)$ denotes the transmission rate, while K_1 and K_2 quantify the reduction in susceptibility due to infection- and vaccine-derived immunity, respectively. The parameters d_0^1 and d_0^2 encode the breadth of cross-immunity (in G.U.) for each immunity source. The second term models the return of recovered individuals to the susceptible class, a process governed by the rate δ and assumed to occur rapidly, since global susceptibility is effectively determined by population-level immunity [26]. We now turn to the infected compartment, which is structured by the phenotype of the infecting virus:

$$\begin{aligned} \frac{\partial I}{\partial t}(x, t) = & \theta(t) \Delta I(x, t) + \beta(x, t) S(t) \frac{I(x, t)}{1 + K_1 \int_X \phi_x(y, d_0^1) M(y, t) dy + K_2 \int_X \phi_x(y, d_0^2) M(y, t) dy} \\ & - \mu I(x, t), \end{aligned} \quad (2)$$

The first term on the right-hand side represents genotype changes within the infected population, arising from mutations that occur during influenza replication inside hosts. Because these mutational processes are assumed independent of the genotype x , the combined effect of many simultaneous host-level mutation events can be modeled as an effective diffusion process in genotype space. Unlike classical neutral drift or selective sweep models, which track the fate of discrete mutations and lineages, our formulation operates at the population level and captures the continuous generation of genetic diversity. Selection is incorporated implicitly through the genotype-dependent transmission rate, which biases the diffusion toward high-fitness regions of the landscape. To incorporate variant plasticity, we introduce a time-dependent diffusion coefficient that scales with the rate of new infections, reflecting the idea that the total number of mutations increases proportionally with viral replication:

$$\theta(t) = \sigma S \int_X \beta(x, t) \frac{I(x, t)}{1 + K_1 \int_X \phi_x(y, d_0^1) M(y, t) dy + K_2 \int_X \phi_x(y, d_0^2) M(y, t) dy} dx.$$

Thus, the diffusion coefficient $\theta(t)$ depends on both the per-infection diffusion rate σ and the total number of new infections. The second term in the infected equation describes new infections arising from susceptible individuals, downregulated by the levels of infection- and vaccine-derived immunity (M and M_v) and their respective efficacies (K_1 and K_2). The third term represents recovery, occurring at rate μ . The effect of population immunity

is included by assuming that the reduction in susceptibility depends on the breadth of immunity. For infection-derived immunity, it is captured through the following expression:

$$\int_X \phi_x(y, d_0^1) M(y, t) dy = \begin{cases} \int_X \frac{M(y, t)}{2} \left(1 + \cos \left(\frac{|y-x|}{d_0^1} \right) \right) & \text{if } |x-y| \leq d_0^1 \\ 0 & \text{elsewhere} \end{cases}$$

An analogous expression is used to represent the reduction in susceptibility mediated by the breadth of vaccine-derived immunity. The cross-immunity kernel was modeled as a cosine-type radial basis function because it provides a smooth, symmetric decline in cross-protection with increasing antigenic distance [29]. It ensures a bounded, continuous function with a natural peak at zero distance and symmetric decay on either side, avoiding artifacts associated with non-smooth kernels. Although no previous study has explicitly employed a cosine formulation, the general approach of using radial basis functions to model cross-immunity has been introduced and explored in earlier work [30]. We then consider the recovered population, which is likewise structured according to the genotype of the virus that caused the infection. Its dynamics are given by the following equation:

$$\frac{\partial R}{\partial t}(x, t) = \mu I(x, t) - \delta R(x, t), \quad (3)$$

The first term on the right-hand side represents the inflow of individuals recovering from infection, while the second captures their return to the susceptible class. In this framework, recovered individuals experience only a brief period of protection before re-entering the susceptible pool [26], since overall susceptibility is determined by the prevailing level of population-immunity. We now describe the evolution of population-immunity, stratified by genotype:

$$\frac{\partial M}{\partial t}(x, t) = k_1 \frac{R(x, t)}{N(1 + K \int_X \phi_x(y, d_0^1) M(y, t) dy + K \int_X \phi_x(y, d_0^2) M_v(y, t) dy)} - \omega_1 M(x, t), \quad (4)$$

The first term on the right-hand side represents the gain in population-immunity generated by new recoveries, while the second accounts for the loss of immunity due to waning. The parameter k denotes the rate at which immunity is upregulated following recovery, and ω specifies the rate of immune decay. We also incorporate a saturation effect in immunity generation that depends on the genotype of pre-existing antibodies. This term reflects the reduced antibody production observed in individuals who are reinfected with the same strain, consistent with immunological memory and B-cell dynamics [31]. We next describe the dynamics of vaccine-induced immunity:

$$\frac{\partial M_v}{\partial t}(x, t) = k_2 \Psi(x, t) - \omega_2 M_v(x, t). \quad (5)$$

Here, the function $\Psi(x, t)$ specifies the distribution of vaccine peptides across the genotype space, representing the strain composition included in the vaccine formulation. It captures different design choices in peptide-based vaccines, characterized by their relative positions and distances from variants 1 and 2. In this work, we examine four peptide-distribution shapes: rectangular, triangular, M-shaped, and Gaussian. Each shape is scaled to have the same total area, ensuring that all formulations correspond to an equal administered antigenic dose. The four vaccine-distribution functions used in our analysis are illustrated in Figure 2. The second term on the right-hand side of the equation corresponds to the waning of vaccine-derived immunity.

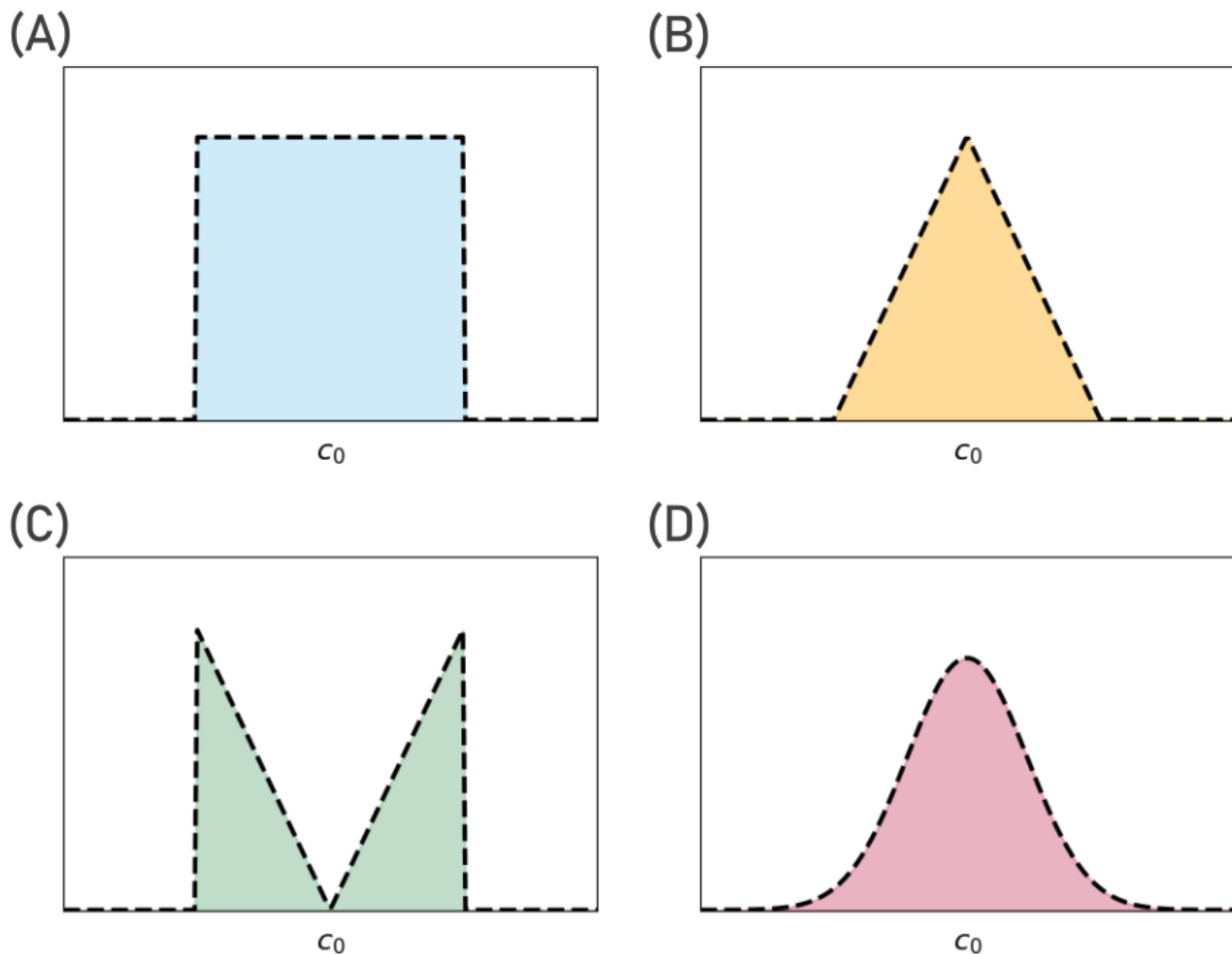


Figure 2. The shapes for the vaccination composition function $\Psi(x, t)$ examined in our analysis: rectangular (A), triangular (B), M-shaped (C), and Gaussian (D).

The simulation is initiated by considering 10 infected individuals with a strain corresponding to variant 1, modeled as a normal distribution with a very small variance equal to 0.02, to account for the relative diversity of strains that consist the initially inhaled virus dose:

$$I(x, 0) = \frac{10}{0.02\sqrt{2\pi}} \exp\left(-\frac{1}{2} \frac{(x-2)^2}{0.02^2}\right) \quad (6)$$

2.2. Numerical Implementation

We parameterize the model using influenza transmission dynamics representative of the United States. Default parameter values are summarized in Table 1. The infection rates for influenza 1 and 2 are obtained by fixing the basic reproduction number and the recovery rate, and then applying the standard expression for the basic reproduction number to derive the infection rate for variant 0:

$$\mathcal{R}_0 = \frac{\beta^0 N}{\mu},$$

where N denotes the total population size.

The model equations are discretized using an explicit Euler scheme, while the diffusion term in Equation (2) is approximated using a second-order central finite-difference operator. Although conditionally stable, these numerical schemes are widely used and provide

consistent and accurate approximations. The temporal and genotype-space discretization steps are set to $dt = 1 \text{ day}^{-1}$ and $dx = 0.25 \text{ G.U.}$, respectively. A consistency analysis for the model in the absence of vaccination examining the impact of the genotype-space discretization is provided in our previous study [22]. All simulations were implemented in Python 3.11. A full run of 1000 days requires approximately 12 s seconds on a workstation equipped with an AMD Threadripper processor and 64 GB of RAM. The source code is available upon reasonable request from the corresponding author. The model is parameterized according to influenza infection. The numerical values of the parameters are given in Table 1. A sensitivity analysis showing the impact of perturbations in the parameters is provided in Appendix A.

Table 1. Default values for the parameters used in the model. G.U. denotes an arbitrary unit in the genotype space.

Parameter	Value	Calibration Method or Reference
N	3.2×10^8	US population
$\mathcal{R}_0$	2.3	Range from 1.8 to 3 [32]
δ	1/10	Short duration [26]
σ	$50 \times 10^{-9} (\text{G.U.})^2$	Arbitrary (drift per each new infection)
μ	1/5	Influenza recovery time is 3–7 days
d_0^1	3 G.U.	3/4 of the distance between the variant centers
d_0^2	variable	Arbitrary
K_1	18	95% protection against transmission [26]
K_2	8	80% protection against transmission [26]
k	3.84	Fitted to seroprevalence data [33]
K	10	Immunity saturation rate [26]
ω_1	$1/(30 \times 6) \text{ day}^{-1}$	Half-life time of infection immunity is 6 months [34]
ω_2	$1/(30 \times 8) \text{ day}^{-1}$	Half-life time of vaccine immunity is 8 months

3. Results

3.1. Variant Emergence in the Absence of Vaccination

We first analyze variant emergence in the absence of vaccination. The epidemic is initialized with 10 infections of variant 1, with variants 1 and 2 separated by a genotypic distance of 3 G.U. and assumed to have identical transmissibility. Over a 1000-day simulation, the epidemic begins with a large wave driven by variant 1, followed by a smaller wave caused by variant 2 due to cross-immunity and a later period of co-circulation of variants 1 and 2 (Figure 3). We calculate the number of infections generated by variants 1 and 2 as follows:

$$I_{new}^1(t) = \int_2^4 \beta(x)S(t) \int_X \frac{I(x,t)}{1 + K_1 \int_X \phi_x(y, d_0^1)M(y,t)dy + K_2 \int_X \phi_x(y, d_0^2)M_v(y,t)dy} dx,$$

$$I_{new}^2(t) = \int_6^7 \beta(x)S(t) \int_X \frac{I(x,t)}{1 + K_1 \int_X \phi_x(y, d_0)M(y,t)dy + K_2 \int_X \phi_x(y, d_0^2)M_v(y,t)dy} dx.$$

Variant 1 is the dominant contributor to transmission, generating 1.02×10^8 infections (58.8%), whereas variant 2 produces 7.13×10^7 infections (41.0%), for a total burden of 1.74×10^8 infections. The rest of the infections were driven by strains not belonging to variants 1 and 2. The wave caused by variant 1 reaches an early peak of 1.06×10^6 daily infections around day 233. Variant 2 peaks later, at 3.88×10^5 daily infections around day 585, producing a smaller secondary wave. These results show that variant 1 dominates early transmission, while variant 2 contributes a delayed but less intense resurgence.

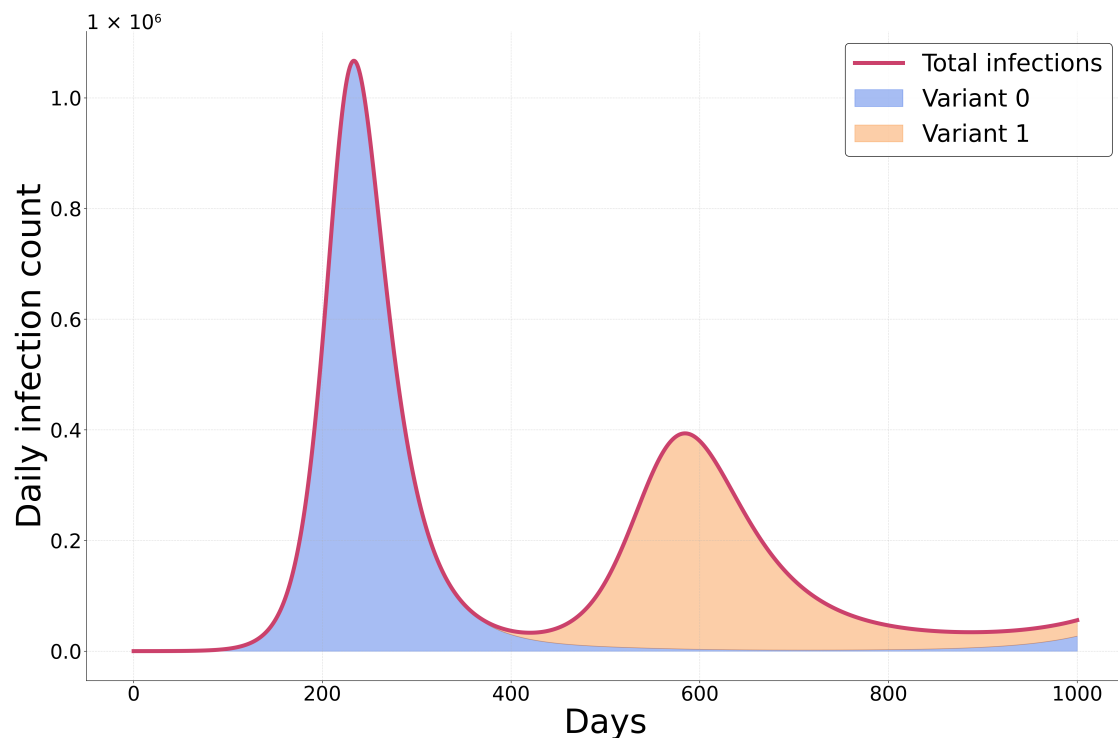


Figure 3. A screenshot of a simulation in the absence of vaccination. It shows an initial wave of infections by variant 1, followed by a milder wave by variant 2. Then, a co-existence between the two variants.

3.2. Narrow-Spectrum Vaccines Minimize the Infection Burden but Promote the Emergence of Other Variants

We introduce vaccination by implementing a campaign beginning on day 450 and ending on day 510, during which 800,000 doses are administered per day, covering a total of 15 million individuals. The choice of this timing and duration was motivated by the tendency to schedule vaccination campaigns in the first weeks of the epidemic and keep them for approximately two months. We first examine a Gaussian vaccine-strain distribution that targets variant 2 and vary its standard deviation while rescaling the amplitude to maintain a constant antigenic dose. Specifically, we explore values of the standard deviation ranging from 0.25 G.U. to 1.5 G.U. Figure 4 presents the resulting epidemic trajectories together with the corresponding genotype distribution of infections. Narrow Gaussian vaccines (small standard deviation values) substantially reduce infections during the variant 2 wave, but this benefit is followed by a resurgence of variant 1. This rebound arises from the weaker cross-immunity generated by variant 2 infections and the limited effectiveness of narrowly targeted vaccines against variant 1.

Using a broad vaccination profile (1.5 G.U.), it remains possible to suppress the emergence of variant 2 while also delaying the subsequent resurgence of variant 1. In the baseline scenario, the total number of infections after day 400 is 75,580,774, with variants 1 and 2 contributing 5.94% and 93.86%, respectively. A narrow vaccine (0.25 G.U.) reduces total infections by 19.12%, but shifts the variant distribution: variant 1 infections increase 3.16-fold, whereas variant 2 infections decrease by 35.96%. Under a broad-spectrum vaccine (1.5 G.U.), total infections decrease by 16.5%, accompanied by a 1.31-fold increase in variant 1 infections and a 19.45% reduction in variant 2 infections. Table 2 summarizes the infection burden for each standard deviation value relative to the baseline (no vaccination).

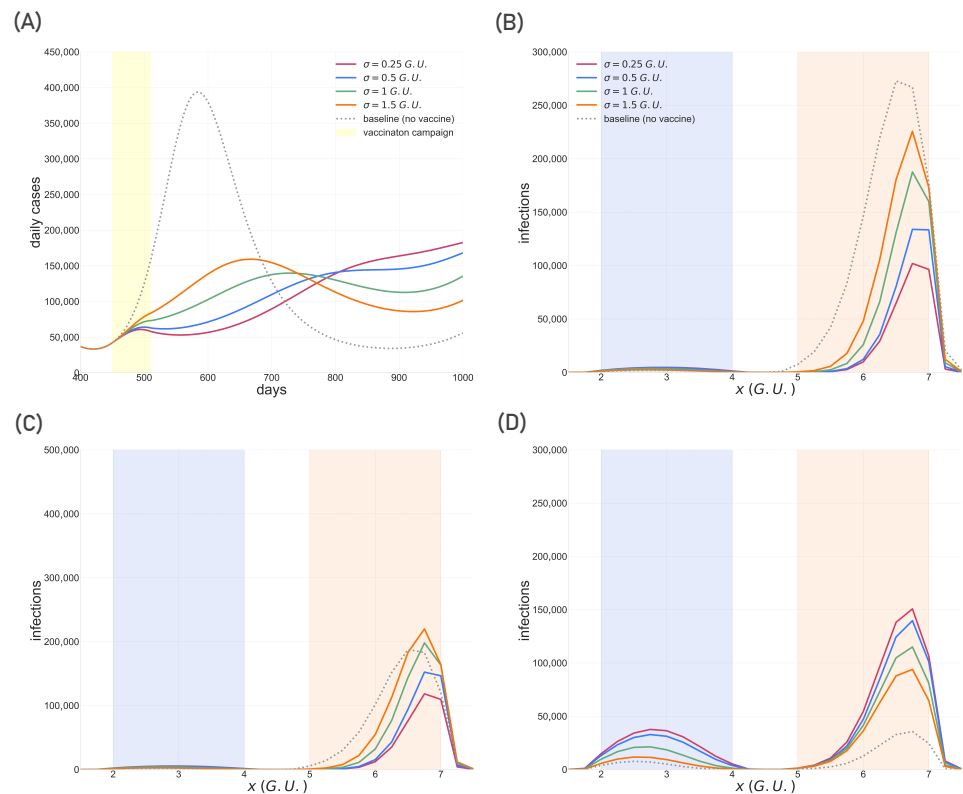


Figure 4. (A) Epidemic curves showing the impact of vaccine distributions on the number of new infections. (B–D) The distributions of infections across the genotype space on days 680, 780, and 900, respectively. The blue and orange ribbons indicate viable spaces for variants 1 and 2, respectively.

Table 2. Relative infections (days 400–1000) compared to baseline (no vaccine) for total, variant 1, and variant 2.

Vaccine Standard Deviation	Total Infections	Variant 1	Variant 2
0.25 G.U.	80.88%	316.25%	66.04%
0.5 G.U.	82.63%	283.74%	69.94%
1 G.U.	83.39%	200.41%	76.04%
1.5 G.U.	83.50%	131.02%	80.55%

3.3. M-Shaped Formulations Are More Effective than Triangular-Shaped for Vaccines with a Reduced Cross-Protection

We extend our analysis by evaluating the effectiveness of additional vaccination-composition shapes: rectangular, triangular, and M-shaped. Our objective is to quantify how changes in vaccine breadth influence the total number of infections under each formulation. When the cross-immunity length of vaccine-elicited immunity is relatively narrow (d_0^2), the M-shaped vaccine substantially suppresses the peak of the variant 1-driven wave and minimizes the overall number of cases (Figure 5A). As the cross-immunity length increases; however, the triangular vaccine formulation becomes more effective, delaying and reducing the magnitude of the variant 2-driven wave (Figure 5B).

To further examine this shift in performance, we compute the number of averted infections across the three vaccine shapes as a function of the breadth of vaccine-induced cross-protection. As shown in Figure 5C, for small cross-immunity breadths, the M-shaped formulation averts the greatest number of infections, followed by the rectangular and triangular formulations. For cross-immunity breadths exceeding 2.4 G.U., this trend reverses: the triangular formulation averts the most infections, followed by the rectangular and M-shaped formulations.

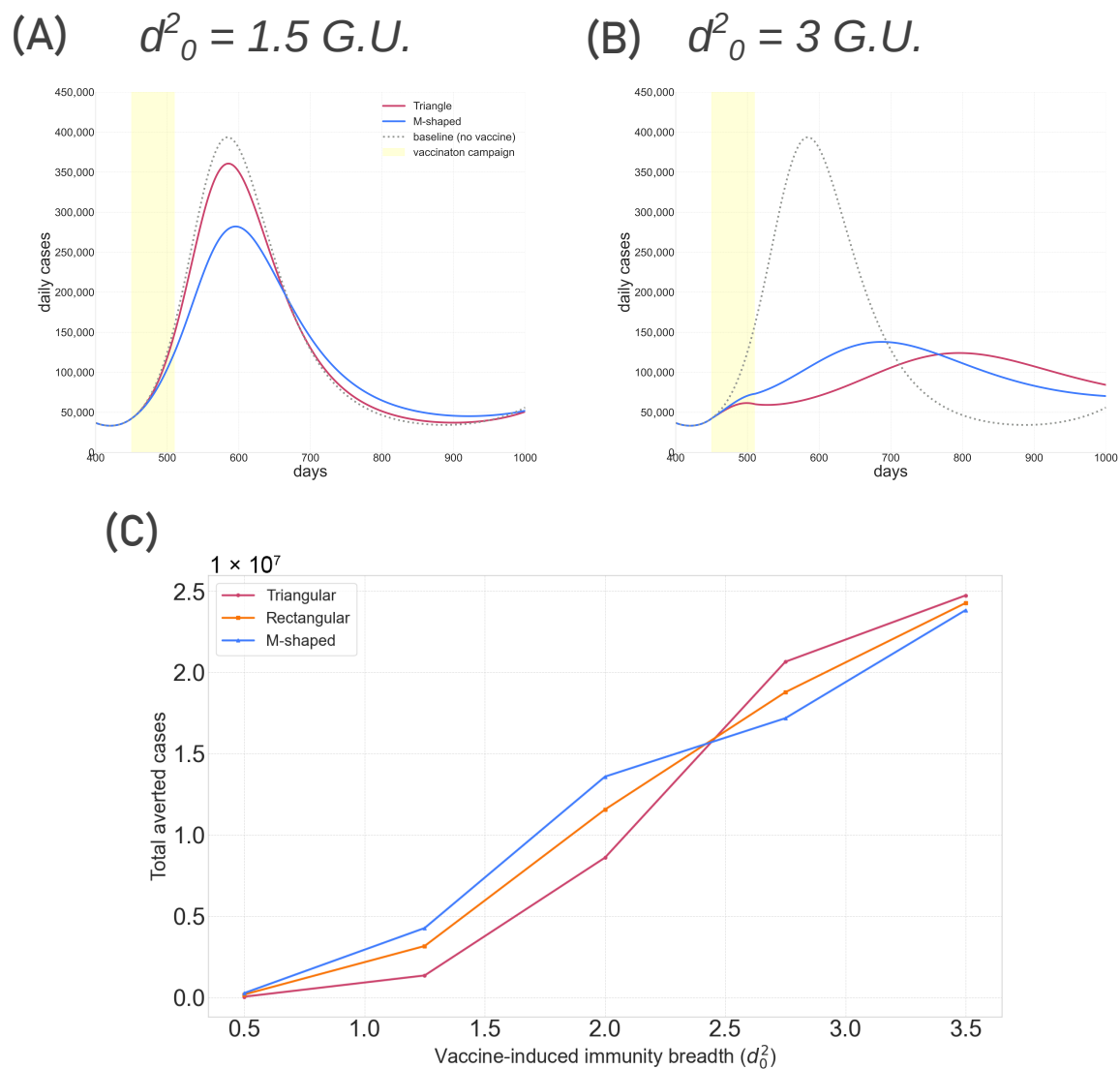


Figure 5. Daily infections following a vaccination campaign using triangular and M-shaped vaccine formulations for two cross-immunity breadths: 1.5 G.U. (A) and 3 G.U. (B). Panel (C) shows the number of averted infections obtained with triangular, rectangular, and M-shaped vaccine compositions across different cross-immunity breadths.

3.4. Broad Cross-Protection Vaccines Avert More Infections When Targeting Inter-Variant Regions

We also examine how the targeted vaccine location c_0 influences the number of averted infections. The vaccine is assumed to target positions ranging from 0.5 G.U. from variant 1 to 3.5 G.U. Our simulations show that when the breadth of vaccine-derived cross-immunity is small ($d_0^2 = 1.5$ G.U.) (Figure 6A), the number of averted infections increases as the targeted location shifts toward variant 2. In this regime, M-shaped vaccines provide greater protection than rectangular or triangular formulations until the target approaches variant 2, beyond which the triangular formulation becomes the most effective. For broader cross-immunity ($d_0^2 = 3$ G.U.) (Figure 6B), the optimal target lies between the two variants: triangular vaccines avert more infections when the target is closer to variant 1, whereas M-shaped vaccines offer greater protection when the target shifts toward variant 2.

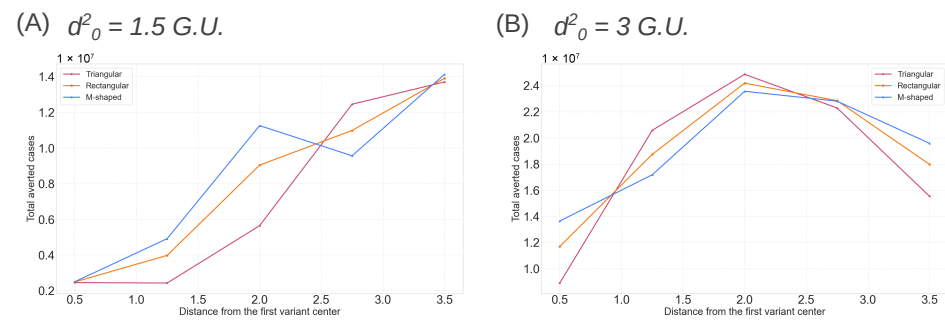


Figure 6. The number of averted deaths as a function of different vaccine distances to the right of variant 1 for lower ($d_0^2 = 1.5$ G.U., (A)) and higher ($d_0^2 = 3$ G.U., (B)) vaccine-induced immunity cross protection.

4. Discussion

This work investigates how vaccine composition influences the emergence of viral variants. We extend our previously developed transmission models, which couple virus evolution with host immunity, by incorporating explicit genotype-dependent vaccine formulations [21,22]. These models were previously parameterized and validated using epidemic data. We first analyze variant emergence in the absence of vaccination. We then perform numerical simulations using a range of vaccine configurations. Focusing initially on a Gaussian formulation, we show that narrow vaccines (smaller standard deviations) reduce the overall infection burden but simultaneously promote antigenic drift, enabling the re-emergence of the original variant. These findings are consistent with prior modeling and immunological studies indicating that conventional, strain-matched vaccines can facilitate variant emergence, whereas broader, cross-protective vaccines help suppress it [20,35,36].

While existing multi-strain models can represent multivalent vaccines by including a finite set of discrete strains or antigenic classes [20], this formulation inherently limits the resolution at which vaccine composition can be explored. In contrast, our framework introduces vaccination as a continuous, genotype-dependent function, allowing fine-grained control over antigenic targeting, breadth, and relative weighting across genotype space. This enables the systematic investigation of subtle design choices, such as targeting intermediate regions between variants or shaping immunity with nonuniform profiles, which are difficult to capture with discrete strain-based models. Our approach therefore generalizes classical multivalent formulations and provides a higher-resolution perspective on how vaccine composition influences evolutionary and epidemiological dynamics.

We then examined additional vaccine-composition functions, including triangular, rectangular, and M-shaped profiles. Our simulations show that M-shaped formulations perform best when vaccine-induced cross-immunity is narrow, whereas triangular formulations avert more infections when cross-protection is broader. This difference likely reflects the ability of triangular profiles to concentrate immunity more effectively around a single genotype. We next investigated the optimal antigenic locations to target in order to limit variant emergence. Our results indicate that vaccines with limited cross-immunity are most effective when centered on the emergent variant, while broadly protective vaccines should target antigenic regions lying between variants. A similar strategy has been employed in the design of “hybrid” influenza vaccines that place antigens in intermediate positions to maximize cross-lineage protection [37].

It is important to note that the triangular and M-shaped vaccine distributions considered in this study are conceptual profiles defined over a one-dimensional genotype (or antigenic) coordinate to probe how antigenic targeting and cross-protection shape variant emergence. Current vaccine manufacturing does not produce continuous antigen distri-

butions of these exact shapes. However, existing multivalent platforms can approximate them by combining a finite set of antigens and tuning their relative doses or display stoichiometries [38]. For example, bimodal (“M-shaped”) targeting can be implemented as bivalent/multivalent mixtures or mosaic nanoparticle co-display of divergent antigens, while triangular weighting can be approximated by selecting multiple antigens across a region and assigning graded component doses [39].

It is important to acknowledge several simplifying assumptions made to facilitate the interpretation of our results. First, the genetic distances in the model were given as a function of an arbitrary unit (G.U.). In practice, it would be possible to link these distances to genetic maps while calibrating the model to actual diseases [2]. Second, the model employs a one-dimensional fitness landscape for the basic reproduction number, consisting of two variants separated by a non-viable region. This representation allows us to capture the antigenic distance between variants, while reducing computational cost and allowing the systemic exploration of theoretical conditions underlying variant emergence. Third, our analysis focuses on a single scenario of variant emergence; in practice, vaccine formulations may influence other evolutionary pathways differently. Fourth, some of the vaccine profiles we consider, such as M-shaped or triangular formulations, or vaccines targeting non-viable antigenic regions, may be challenging to realize experimentally. We nevertheless include them to evaluate their theoretical performance and to identify design principles that could inform future vaccine development. Finally, we adopted a uniform vaccination window to reduce complexity and isolate the effects of vaccine composition on variant emergence. Incorporating dynamic rollout, boosters, or age stratification would introduce additional confounding factors that obscure the mechanistic role of formulation shape and breadth. These extensions are important but fall outside the scope of the present work and will be explored in future studies.

5. Conclusions

Overall, our work highlights the central role of vaccine-induced cross-protection in mitigating the impact of variant emergence and provides design principles to guide the development and deployment of both narrow and broad vaccine formulations. It is important to consider the disease and population features when considering these principles. Although the present study is purely theoretical, future work will apply this framework to specific pathogens—such as influenza—by incorporating pathogen-specific transmission characteristics and integrating within-host relapse dynamics through extensions of our genotype-structured viral kinetic model [25]. Future extensions will also incorporate severity-stratified states to more directly link vaccine formulation to clinical outcomes.

Author Contributions: Conceptualization, A.B. and B.D.-R.; methodology, A.B. and B.D.-R.; software, A.B. and B.D.-R.; validation, A.B. and B.D.-R., A.B. and B.D.-R.; formal analysis, A.B. and B.D.-R.; investigation, A.B. and B.D.-R.; writing—original draft preparation, A.B. and B.D.-R.; writing—review and editing, A.B. and B.D.-R.; visualization, A.B. and B.D.-R.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: All equations, parameter values, and numerical methods necessary to reproduce this work are disclosed. The author is available upon reasonable request to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A. Sensitivity Analysis

To evaluate the influence of model parameters on our results, we performed a sensitivity analysis by comparing the relative reduction in cases across different vaccine formulations while slightly varying key parameters. We computed the case reduction relative to the no-vaccination scenario for the case when $d_0^2 = 3$ G.U. Across all parameter values, the ranking of vaccine formulations remained unchanged with only minor quantitative variations (Figure A1). The triangular formulation yielded maximal reduction, followed by the Gaussian (with a standard deviation equal 0.5 G.U.), the uniform, and then the M-shaped.

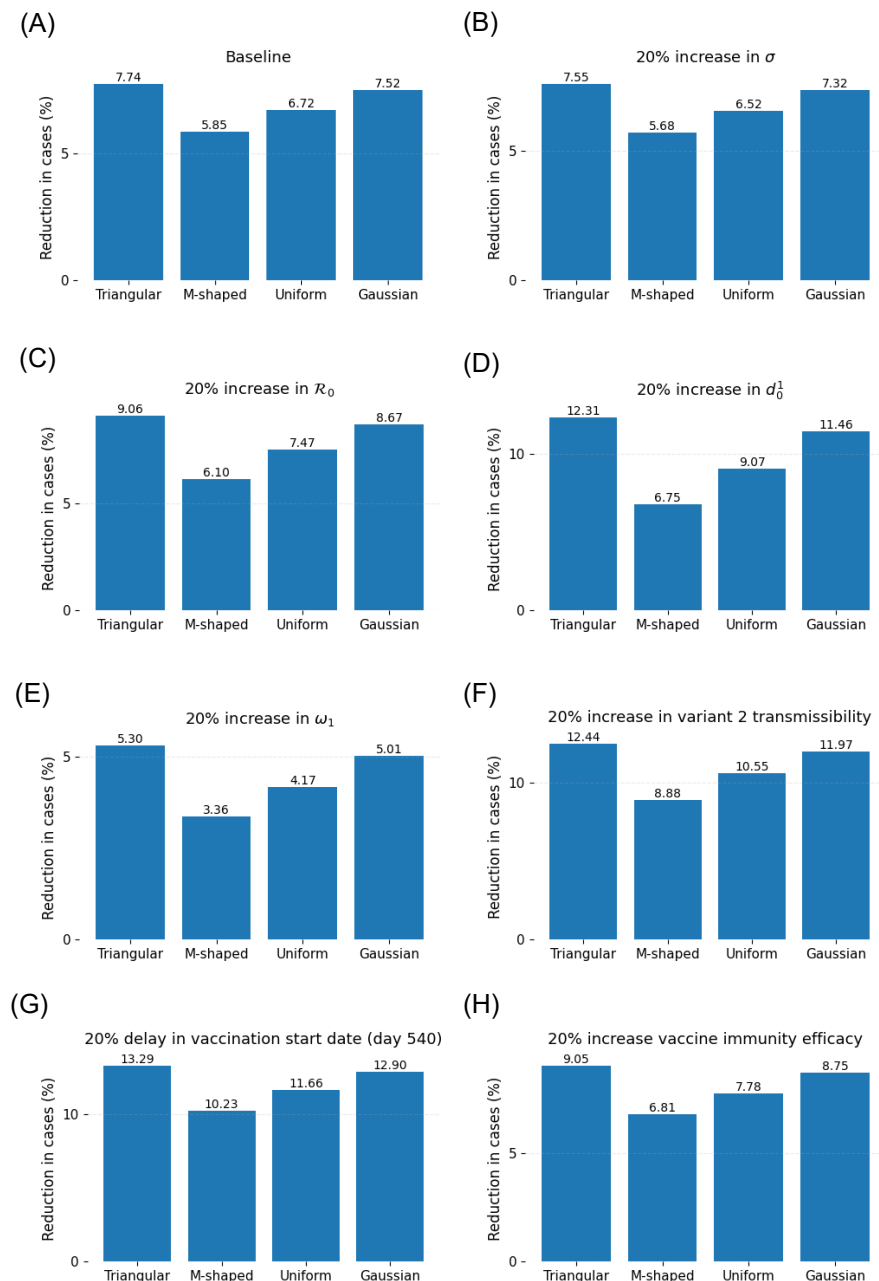


Figure A1. Results of the sensitivity analysis, showing how variations in model parameters affect the reduction in case numbers. It includes a baseline case (A), as well as results obtained when slightly changing σ (B), R_0 (C), d_0^1 (D), ω_1 (E), variant 2 transmissibility (F), vaccination campaign starting date (G), and vaccine efficacy (H).

References

- Smith, D.J.; Lapedes, A.S.; Forrest, S.; de Jong, J.C.; Osterhaus, A.D.; Fouchier, R.A.; Cox, N.J.; Perelson, A.S. Modeling the effects of updating the influenza vaccine on the efficacy of repeated vaccination. In *International Congress Series*; Elsevier: Amsterdam, The Netherlands, 2001; Volume 1219, pp. 655–660.
- Smith, D.J.; Lapedes, A.S.; De Jong, J.C.; Bestebroer, T.M.; Rimmelzwaan, G.F.; Osterhaus, A.D.; Fouchier, R.A. Mapping the antigenic and genetic evolution of influenza virus. *Science* **2004**, *305*, 371–376. [\[CrossRef\]](#)
- Katz, J.M.; Hancock, K.; Xu, X. Serologic assays for influenza surveillance, diagnosis and vaccine evaluation. *Expert Rev. Anti-Infect. Ther.* **2011**, *9*, 669–683. [\[CrossRef\]](#) [\[PubMed\]](#)
- Influenza, A. Recommended composition of influenza virus vaccines for use in the 2013–2014 northern hemisphere influenza season. *Wkly. Epidemiol. Rec.* **2013**, *8*, 101–114.
- Jackson, M.L.; Chung, J.R.; Jackson, L.A.; Phillips, C.H.; Benoit, J.; Monto, A.S.; Martin, E.T.; Belongia, E.A.; McLean, H.Q.; Gaglani, M.; et al. Influenza vaccine effectiveness in the United States during the 2015–2016 season. *N. Engl. J. Med.* **2017**, *377*, 534–543. [\[CrossRef\]](#) [\[PubMed\]](#)
- Rueckert, C.; Guzmán, C.A. Vaccines: From empirical development to rational design. *PLoS Pathog.* **2012**, *8*, e1003001. [\[CrossRef\]](#)
- Neher, R.; Bedford, T. Prediction, dynamics, and visualization of antigenic phenotypes of seasonal influenza viruses. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, E1701–E1709. [\[CrossRef\]](#)
- Cobey, S. Pathogen evolution and the immunological niche. *Nat. Rev. Immunol.* **2018**, *18*, 231–240. [\[CrossRef\]](#)
- Nachbagauer, R.; Palese, P. Is a universal influenza virus vaccine possible? *Annu. Rev. Med.* **2020**, *71*, 315–327. [\[CrossRef\]](#)
- Huber, V.C.; Thomas, P.G.; McCullers, J.A. A multi-valent vaccine approach that elicits broad immunity within an influenza subtype. *Vaccine* **2009**, *27*, 1192–1200. [\[CrossRef\]](#)
- Wilding, K.M.; Molina-París, C.; Kubicek-Sutherland, J.Z.; McMahon, B.; Perelson, A.S.; Ribeiro, R.M. A consensus mathematical model of vaccine-induced antibody dynamics for multiple vaccine platforms and pathogens. *Front. Immunol.* **2025**, *16*, 1596518. [\[CrossRef\]](#)
- Xu, Z.; Song, J.; Zhang, H.; Wei, Z.; Wei, D.; Yang, G.; Demongeot, J.; Zeng, Q. A mathematical model simulating the adaptive immune response in various vaccines and vaccination strategies. *Sci. Rep.* **2024**, *14*, 23995. [\[CrossRef\]](#) [\[PubMed\]](#)
- Leon, C.; Tokarev, A.; Bouchnita, A.; Volpert, V. Modelling of the innate and adaptive immune response to SARS viral infection, cytokine storm and vaccination. *Vaccines* **2023**, *11*, 127. [\[CrossRef\]](#) [\[PubMed\]](#)
- Béraud, G. Mathematical models and vaccination strategies. *Vaccine* **2018**, *36*, 5366–5372. [\[CrossRef\]](#) [\[PubMed\]](#)
- Shankar, M.; Hartner, A.M.; Arnold, C.R.; Gayawan, E.; Kang, H.; Kim, J.H.; Gilani, G.N.; Cori, A.; Fu, H.; Jit, M.; et al. Mathematical modelling to inform outbreak response vaccination. *arXiv* **2024**, arXiv:2410.13923. [\[CrossRef\]](#)
- Jung, S.m.; Loo, S.L.; Howerton, E.; Contamin, L.; Smith, C.P.; Carcelén, E.C.; Yan, K.; Bents, S.J.; Levander, J.; Espino, J.; et al. Potential impact of annual vaccination with reformulated COVID-19 vaccines: Lessons from the US COVID-19 scenario modeling hub. *PLoS Med.* **2024**, *21*, e1004387. [\[CrossRef\]](#)
- Li, Q.; Huang, Y. Optimizing global COVID-19 vaccine allocation: An agent-based computational model of 148 countries. *PLoS Comput. Biol.* **2022**, *18*, e1010463. [\[CrossRef\]](#)
- Andreu-Villarroya, C.; Villanueva, R.J.; González-Parra, G. Mathematical modeling for estimating influenza vaccine efficacy: A case study of the Valencian Community, Spain. *Infect. Dis. Model.* **2024**, *9*, 744–762. [\[CrossRef\]](#)
- Otunuga, O.M. Analysis of multi-strain infection of vaccinated and recovered population through epidemic model: Application to COVID-19. *PLoS ONE* **2022**, *17*, e0271446. [\[CrossRef\]](#)
- Subramanian, R.; Graham, A.L.; Grenfell, B.T.; Arinaminpathy, N. Universal or specific? A modeling-based comparison of broad-spectrum influenza vaccines against conventional, strain-matched vaccines. *PLoS Comput. Biol.* **2016**, *12*, e1005204. [\[CrossRef\]](#)
- Bouchnita, A.; Bandekar, S.R.; Bi, K.; Rouhani, B.D.; Fox, S.J.; Garcia, J.A. The interplay between evolutionary and immunological dynamics regulates virus variant emergence and competition. *Math. Model. Nat. Phenom.* **2025**, *20*, 14. [\[CrossRef\]](#)
- Bouchnita, A.; Djafari-Rouhani, B. Integrating genomic, climatic, and immunological factors to analyze seasonal patterns of influenza variants. *Symmetry* **2024**, *16*, 943. [\[CrossRef\]](#)
- Bouchnita, A.; Volpert, V. Phenotype-structured model of intra-clonal heterogeneity and drug resistance in multiple myeloma. *J. Theor. Biol.* **2024**, *576*, 111652. [\[CrossRef\]](#) [\[PubMed\]](#)
- Bessonov, N.; Reinberg, N.; Volpert, V. Mathematics of Darwin’s diagram. *Math. Model. Nat. Phenom.* **2014**, *9*, 5–25. [\[CrossRef\]](#)
- Bouchnita, A. Genotype-structured modeling of variant emergence and its impact on virus infection. *Mathematics* **2025**, *13*, 167. [\[CrossRef\]](#)
- Bouchnita, A.; Bi, K.; Fox, S.J.; Meyers, L.A. Projecting Omicron scenarios in the US while tracking population-level immunity. *Epidemics* **2024**, *46*, 100746. [\[CrossRef\]](#)
- Bi, K.; Bandekar, S.R.; Bouchnita, A.; Fox, S.J.; Meyers, L.A. Annual Hospitalizations for COVID-19, Influenza, and Respiratory Syncytial Virus, United States, 2023–2024. *Emerg. Infect. Dis.* **2025**, *31*, 636. [\[CrossRef\]](#)

28. Banerjee, M.; Lipniacki, T.; d’Onofrio, A.; Volpert, V. Epidemic model with strain-dependent transmission rate. *Commun. Nonlinear Sci. Numer. Simul.* **2022**, *114*, 106641. [[CrossRef](#)]
29. Luo, G.; Yang, Z.; Zhan, C.; Zhang, Q. Identification of nonlinear dynamical system based on raised-cosine radial basis function neural networks. *Neural Process. Lett.* **2021**, *53*, 355–374. [[CrossRef](#)]
30. Adams, B.; Sasaki, A. Cross-immunity, invasion and coexistence of pathogen strains in epidemiological models with one-dimensional antigenic space. *Math. Biosci.* **2007**, *210*, 680–699. [[CrossRef](#)]
31. Prechl, J. A generalized quantitative antibody homeostasis model: Antigen saturation, natural antibodies and a quantitative antibody network. *Clin. Transl. Immunol.* **2017**, *6*, e131. [[CrossRef](#)]
32. Nikbakht, R.; Baneshi, M.R.; Bahrampour, A. Estimation of the basic reproduction number and vaccination coverage of influenza in the United States (2017–18). *J. Res. Health Sci.* **2018**, *18*, e00427.
33. Suntronwong, N.; Vichaiwattana, P.; Wongsrisang, L.; Klinfueng, S.; Korkong, S.; Thongmee, T.; Wanlapakorn, N.; Poovorawan, Y. Prevalence of antibodies against seasonal influenza A and B viruses among older adults in rural Thailand: A cross-sectional study. *PLoS ONE* **2021**, *16*, e0256475. [[CrossRef](#)] [[PubMed](#)]
34. Skowronski, D.M.; Tweed, S.A.; Tweed, S.A.; De Serres, G. Rapid decline of influenza vaccine—Induced antibody in the elderly: Is it real, or is it relevant? *J. Infect. Dis.* **2008**, *197*, 490–502. [[CrossRef](#)] [[PubMed](#)]
35. Boni, M.F. Vaccination and antigenic drift in influenza. *Vaccine* **2008**, *26*, C8–C14. [[CrossRef](#)] [[PubMed](#)]
36. Meijers, M.; Vanshylla, K.; Gruell, H.; Klein, F.; Lässig, M. Predicting in vivo escape dynamics of HIV-1 from a broadly neutralizing antibody. *Proc. Natl. Acad. Sci. USA* **2021**, *118*, e2104651118. [[CrossRef](#)]
37. Gu, C.; Babujee, L.; Pattinson, D.; Chiba, S.; Jester, P.; Maemura, T.; Neumann, G.; Kawaoka, Y. Development of broadly protective influenza B vaccines. *NPJ Vaccines* **2025**, *10*, 2. [[CrossRef](#)]
38. Hernandez-Davies, J.E.; Felgner, J.; Strohmeier, S.; Pone, E.J.; Jain, A.; Jan, S.; Nakajima, R.; Jasinskas, A.; Strahsburger, E.; Krammer, F.; et al. Administration of multivalent influenza virus recombinant hemagglutinin vaccine in combination-adjuvant elicits broad reactivity beyond the vaccine components. *Front. Immunol.* **2021**, *12*, 692151. [[CrossRef](#)]
39. Cohen, A.A.; van Doremalen, N.; Greaney, A.J.; Andersen, H.; Sharma, A.; Starr, T.N.; Keeffe, J.R.; Fan, C.; Schulz, J.E.; Gnanaprasagam, P.N.; et al. Mosaic RBD nanoparticles protect against challenge by diverse sarbecoviruses in animal models. *Science* **2022**, *377*, eabq0839. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.