Modeling Healthcare Policy: From Calibration to Optimization

Serin Lee

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington

2024

Committee:

Shan Liu, Chair

Zelda B. Zabinsky, Chair

Youngjun Choe

Program Authorized to Offer Degree:
Industrial and Systems Engineering

University of Washington

Abstract

Modeling Healthcare Policy: From Calibration to Optimization

Serin Lee

Co-Chairs of the Supervisory Committee:
Shan Liu
Industrial and Systems Engineering

Zelda B. Zabinsky
Industrial and Systems Engineering

Setting effective healthcare policy is complex, particularly due to the dynamic and heterogeneous nature of individual behaviors. While modeling studies offer valuable insights, their computational complexity and calibration requirements can limit practical applications. An important question when identifying optimal healthcare policies is how to consider heterogeneous individual health behavior dynamics while at the same time find ways to consume less time and computational resources.

This dissertation addresses three main objectives. The first objective aims to optimize public health interventions to minimize the disease burden of the COVID-19 pandemic. An agent-based model (ABM) is developed to evaluate non-pharmaceutical interventions (NPIs), and vaccination policies under continuous virus mutation. This ABM simulates a heterogeneous population and offers flexibility for addressing diverse policy questions. By addressing parameter uncertainty through calibration and simulating multiple scenarios, the model identifies robust strategies, such as periodic vaccination and adaptive social distancing, for effective disease control.

The second objective is to design optimal vaccination promotion campaigns that increase vaccination uptake and improve public health. This approach integrates coupled dynamics, social contagion, and evolutionary game theory to model how vaccination behavior shifts within an ongoing epidemic and with word-of-mouth vaccination campaigns. This model overcomes the limitations of prior studies that assumed static vaccination willingness in vaccination allocation problem. The study offers population-level insights into how resources and messaging should be targeted across

demographic groups while considering the societal contexts surrounding vaccination within communities.

The final objective is to enhance calibration practices for simulation models by introducing a representative calibration framework. This approach is particularly valuable for complex models where uncertainties exist, and evaluation for policy analysis is computationally expensive. This framework identifies minimal sets of parameter values to limit computational expense while capturing diverse model behaviors. By focusing on representativeness in calibration, this research yields reliable implications for real-world decision-making, filling a gap where other methods emphasize precision at the risk of not accounting for data and model uncertainty.

Overall, this dissertation advances healthcare policy modeling by addressing complex and heterogeneous individual behaviors, as well as addressing uncertainties arising from data limitations and simulation complexity through calibration. This research will equip healthcare policymakers to derive informed, data-driven insights from modeling studies, despite model uncertainties and complexities.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The past five years at the University of Washington have been an unforgettable and amazing experience, both academically and personally. I would like to express my gratitude to everyone who has supported me on this journey—my advisors, mentors, family, and friends. Your guidance and encouragement were invaluable in my completing this dissertation.

First and foremost, I would like to express my deepest gratitude to Dr. Shan Liu and Dr. Zelda B. Zabinsky. They taught me how to approach research with curiosity and a mindset, starting with simpler problems to formulate a comprehensive understanding. These invaluable lessons will undoubtedly guide me throughout my career as I continue to grow as a scientist. Their generosity, patience, warmth, and kindness in treating me as a person have taught me what it means to be a good person.

I would also like to extend my sincere thanks to my committee members, Dr. Youngjung Choe and Dr. Emma S. Spiro, for their insightful comments and feedback. Their suggestions inspired me to explore new questions in my research, broadening the scope and impact of my work. Additionally, I am deeply grateful to my collaborators over the years, including Dr. Jennifer M. Ross, Dr. Judith N. Wasserheit, Dr. Jamie Cohen, Dr. Daniel J. Klein, Dr. Christopher Weyant, Dr. Jeremy Goldhaber-Fiebert, and many others. You all have taught me about disease and healthcare policy modeling. This dissertation would not have been possible without your invaluable contributions.

My ISE colleagues have been truly supportive, offering both friendship and professional guidance over these five years. Special thanks to Chris, Pete, Xiaonan, Dr. Chelsea, Dr. Ameer, Dr. Yilun, Dr. William, Sami, Cherin, Danielle, and Yinsheng for always being there and helping me overcome every obstacle, big and small, throughout this PhD journey.

I have been very fortunate to meet exception Korean friends in Seattle. I would like to extend special thanks to Dr. Jeehyun and Dr. Gounah for being dedicated work buddies, Dr. Doori and Dr. Yunjeong for being close friends with whom I could discuss any problem, and my tennis buddies and Sunsamo members, Sunho, Jaehoon, Yunbi, Kihyuk, Steven, and Yungeun, for spending

countless nights together, providing enjoyment and relief throughout graduate life.

I extend my heartfelt gratitude to my mentor, Dr. Soohyun, who encouraged me to pursue my doctoral studies abroad, even when I doubted my own abilities. Our email exchanges, which I will always cherish, gave me strength and motivation throughout this journey. I am also deeply grateful to my lifelong friends from high school, Yongbin and Esther, who have been by my side through every twist and turn of life's path.

Last but not least, I extend my utmost appreciation to my family. My parents, Heejung and Yeonhuh, my siblings, Sehee and Byeongchoel, and my grandmother Chungja, have provided me with unconditional love, understanding, and support throughout my entire life. Thank you for providing a safe haven whenever I faced difficulties or felt stuck.

Chapter 1

INTRODUCTION

1.1 Overview

Healthcare policies must be carefully designed to consider individuals' heterogeneous and dynamic responses. An urgent research question is how to identify optimal healthcare policies that take into account the complex and dynamic individual behaviors for infectious diseases.

Modeling studies provide valuable insights to policymakers, particularly in unprecedented public health crises like the COVID-19 pandemic. From projecting epidemic trajectories to evaluating the effectiveness of interventions, compartmental and agent-based simulation models (ABMs) [3, 16, 76, 86, 87, 90, 110] informed crucial choices regarding mask mandates, social distancing measures, and vaccine prioritization policy. These modeling studies have played a pivotal role in informing policymakers' decisions during controversies surrounding public health measures.

Specifically, ABMs are powerful simulation tools to understand and predict the outcomes of policies in complex systems. They are particularly well-suited for health policy analysis because they can model heterogeneous individual behaviors, dynamic decision-making, and the interactions of individuals ("agents") within realistic environments. These strengths allow ABMs to be applied in various fields, including urban traffic simulation [99, 109], economic decision-making [130], and rumor spreading [38].

However, ABMs present challenges. Their reliance on Markov Chain Monte Carlo (MCMC)-type simulations introduces randomness, requiring multiple replications that can be computationally demanding. Calibration to fit unknown parameters to observed data and validate the model adds further complexity, potentially introducing uncertainty due to data limitations or an overly flexible model structure. Appropriately addressing parameter uncertainty is crucial for understanding how it impacts the clarity and reliability of policy insights derived from ABMs [3, 90, 110].

This dissertation presents two ABM studies that identify optimal COVID-19 policies, simulating non-pharmaceutical interventions (NPIs) [86], and incorporating vaccination policies under constant viral mutations [87]. This research contributes to the development of a flexible ABM

framework adaptable to diverse policy questions. Parameter uncertainty is addressed by identifying representative parameter sets for calibration and simulates a range of possible mutation scenarios. With this approach, I find that mask wearing, adaptive social distancing, and periodic vaccination are essential for robust COVID-19 control under diverse real-world scenarios.

While previous research highlights the need for periodic vaccination to control the COVID-19 pandemic, low vaccination willingness remains a constant challenge. Modeling vaccination behavior is complex due to its dynamic nature, as individuals' decisions change in response to evolving information about vaccine safety and effectiveness [71], disease severity, and perceived risks and benefits [32]. These perceptions are further impacted by factors like misinformation [94] and government campaigns [147]. Understanding and incorporating these dynamics into models is crucial for designing effective campaigns.

Existing research on dynamic health-related behaviors employ approaches like coupled disease-opinion dynamics [42, 95], social contagion [13, 68, 151], game theory [13, 151]. However, these models often overlook the optimal resource allocation to maximize public health benefits. Current allocation research primarily uses operations research techniques to address resource limitations, often assuming static vaccination willingness [73, 89, 128]. This highlights the need for developing optimal vaccination campaigns that address the dynamic and varied vaccine willingness. While other fields, like political campaigns and marketing [6, 45, 69], use resource allocation optimization models to influence behavior, they often lack the dynamic elements and focus on simplified network structures. This limits their applicability to health policy problem, where individual characteristics such as demographics plays a crucial role in outcomes.

To address this research gap, this dissertation models the complex interplay between disease transmission, dynamic vaccination behavior, and targeted vaccination campaigns across demographic groups. With the use of a networked compartmental model, this research contributes to identify simple yet intuitive optimal vaccination campaign strategies. Findings highlight the context-dependent nature of vaccination campaign effectiveness and allocation strategies, emphasizing that optimal regional and age-based targeting should depend on societal vaccination characteristics.

Finally, model calibration, which is the process of fitting unknown model parameters to observed data, is a challenge in both ABMs and compartmental models. Fitting a model to perfectly match data, especially data with inherent uncertainty or error, can result in the model becoming overly tailored to a specific dataset and losing its ability to generalize to unseen data, making it

unreliable to represent real-world scenarios. To address this issue, there is a need for representative calibration practices that capture a wide range of uncertainties, allowing models to better reflect the variability inherent in real-world healthcare systems and enhance their generalizability.

Various calibration algorithms exist, including traditional methods like Latin hypercube sampling [102] and more advanced techniques employing black-box optimization [34, 40] and Bayesian optimization approaches [2, 63, 66]. While these methods prioritize computational efficiency and accuracy in finding calibration parameters, they often overlook the importance of robustness and representativeness of calibration results. Bayesian methods offer insight into parameter uncertainty, but in policy analysis, the common practice of selecting top parameter sets based on data matching score can limit representativeness. These calibration approaches can lead either to a small number of optimal parameter sets that insufficiently reflect model variability, or a large number of feasible sets that is computationally burdensome for policy analysis.

The final study of this dissertation tackles this problem by introducing a representative calibration framework. This approach aims to identify a minimal set of parameter sets that effectively capture a wide range of potential model behavior. This approach is particularly valuable for complex models where simulations are computationally expensive yet uncertainties remain high due to data limitations. By ensuring diverse calibrations, models are empowered with greater reliability and robustness for real-world decision-making.

1.2 Research Objective and Approach

The goal of this dissertation is to solve critical public health problems by applying and advancing healthcare research approaches, specifically considering complex dynamic and heterogeneous behavior and data uncertainties. To achieve this goal, I address the following three objectives:

(1) *Design optimal public health interventions to minimize disease burden of COVID-19 pandemic,*

(2) *Optimize vaccination campaigns to increase vaccination uptake and improve public health,*

(3) *Develop a calibration approach for complex simulation models that prioritizes representativeness to account for uncertainty.*

In Chapter 2, the first objective is met by identifying optimal public health policies that minimize the disease burden of COVID-19. A detailed agent-based simulation is developed that models

the spread of the virus through a synthetic population network in King County, WA. The model is calibrated against 2020 epidemic data and is used to simulate two scenarios. The first scenario explores the effects of NPIs and determine effective policies. These NPIs include social distancing, mask usage, school closures, and testing and contact tracing. The second scenario incorporates vaccination policies while considering the viral characteristics of SARS-CoV-2. Additionally, it explores factors like immunity duration and vaccine effectiveness against emerging variants. The goal is to identify long-term strategies for pandemic control.

This research offers several contributions. First, it develops a flexible ABM framework that can be adapted to answer diverse policy questions. Unlike traditional methods that indirectly estimate policy impact through disease parameter modification, this framework directly models specific public health interventions. This allows for explicit analysis of the relationship between policies and disease outcomes. Second, this model directly addresses parameter uncertainty and explains its impact on policy outcomes. For example, we identify two parameter sets that demonstrate different scenarios regarding daily interaction patterns, one with high interaction rates in households, neighborhoods, and workplaces, and another with high interaction rates in schools. In both scenarios, social distancing and face mask use consistently reduce disease burden, whereas school closure is only effective in the latter case. Finally, this model was instrumental in identifying effective policies under the unique challenges of COVID-19. The results of this research were shared with local public health leaders and modeling teams at the Washington State Department of Health, directly informing policy decisions.

In Chapter 3, the second objective is met by designing targeted vaccination campaigns that improve health outcomes. A model is designed that combines coupled dynamics, evolutionary game theory, and social networks analysis to explore the complex relationship between disease transmission, changing vaccination opinions, and word-of-mouth campaigns. This approach reveals that societal factors, including perceived vaccine risk and whether individuals prioritize logic or emotion, significantly impact campaign effectiveness and optimal allocation.

This research addresses a research gap in vaccination campaign design by developing a comprehensive networked compartmental model that incorporates dynamic vaccination behavior and coupled dynamics to optimize vaccination campaigns. This approach enables the consideration of changing individual behaviors for designing more effective and targeted vaccination campaign allocations. Results highlight the context-dependent nature of vaccination campaigns, demonstrating the importance of allocating vaccine resources differently across demographic groups based

on societal characteristics. This work has significant implications in the current context of vaccine hesitancy, where tailored and strategic approaches are more crucial than ever.

In Chapter 4, the third objective is satisfied by developing a calibration approach for simulation models that captures both model and data uncertainties, while being computationally efficient. This approach has two-stages, designed to identify a set of parameter values whose model outcomes closely match target data but also effectively capture the inherent uncertainty within the model. The first stage is to apply a black-box optimization algorithm to minimize a loss function that includes the predicted outcomes and data outcomes as arguments. This results parameter sets that represent acceptable predicted outcomes. The second stage makes use of clustering methods to identify a minimal subset of parameter sets that maximizes diversity for future analysis.

This framework offers a comprehensive calibration approach, enhancing model reliability and robustness for complex problems with limited data and variable model behavior. Unlike many existing methods that focus solely on speed and precision, this framework prioritizes representativeness to capture diverse model behaviors with computational efficiency. This ensures the model reflects the full spectrum of potential outcomes with minimal efforts, addressing a critical limitation in models with high uncertainty due to data limitations and model flexibility.

Finally, Chapter 5 concludes the dissertation and provides directions for future research.

Chapter 2

OPTIMAL PUBLIC HEALTH INTERVENTIONS TO MINIMIZE DISEASE BURDEN OF COVID-19 PANDEMIC

2.1 Introduction

Since the beginning of the COVID-19 pandemic, policymakers have been striving to balance public health interventions to keep people safe and minimize social disruptions. In this chapter, we address research objective (1) by identifying optimal public health interventions on the control of COVID-19 pandemic in two situations. In the first study, we explore the effects of non-pharmaceutical interventions (NPIs) without vaccination during the early stage of the pandemic. The study has been published in Medical Decision Making [86]. In the second study, we explore the effects of periodical vaccination and NPIs. In the study, we consider complex known uncertainties about SARS-CoV-2 such as individuals' heterogeneous immune responses to natural infections and vaccination, and the emergence of more contagious or serious disease-causing viral variants. The second study has been published in AJPM Focus [87].

To conduct this research, a large-scale agent-based model (ABM) is developed and calibrated to COVID-19 outbreaks in the greater Seattle area, simulating the behavior of approximately 1.9 million individuals and their responses to policies based on demographics and health status. This research contributes to the development of an adaptable ABM framework that can be effectively applied to inform diverse policy decisions. This research addresses parameter uncertainty by identifying representative calibration parameter sets and simulating a range of possible mutation scenarios. This approach highlights the importance of periodic vaccination and adaptive social distancing as keys to robust COVID-19 control in complex real-world environments.

This chapter is organized as follows. Section 2.2 contains the first study that identifies the impact of NPIs on reopening the society. Section 2.3 contains the second study that identifies the impact of NPIs and vaccinations with viral variants and waning immunity. Finally, Section 2.4 summarises the conclusion and discussion of this chapter.

2.2 Impact of Nonpharmaceutical Interventions on Reopening Society

2.2.1 Introduction

The COVID-19 pandemic caused many deaths, economic losses, and disruptions to American society. As many parts of the United States effectively flattened the epidemic curve to control the first wave of outbreaks and reduce hospital overflow through a combination of nonpharmaceutical interventions (NPIs) in the spring and summer of 2020, stringent government policies such as lockdowns incurred significant societal costs. The federal and many state and local governments have been striving to ease restrictions while implementing other interventions such as expanded testing, contact tracing, and face mask use. Unfortunately, many states experienced second or third waves of infection after relaxing restrictions in the fall of 2020.

Modeling studies can provide policymakers with evidence to inform politically sensitive decisions amid large uncertainty about SARS-CoV-2 transmission in the community. Forecasting models and mechanistic models are the two main modeling approaches. The former focuses on prediction in the short-term using statistical methods and fitting historical data, whereas the latter focuses on simulating plausible scenarios of infection spread under various assumptions to account for nonlinear feedback and project long-term outcomes [61]. The ISPOR-SMDM Modeling Good Research Practices Task Force provides a comprehensive overview and recommendations for different modeling approaches [21]. A more detailed discussion of different types of models for dynamic transmission can be found in Pitman et al [113]. The Centers for Disease Control and Prevention (CDC) reports weekly national forecasts of new COVID-19 cases, hospitalizations, and deaths from more than 35 mathematical and statistical models [22]. A number of mathematical models have measured the effectiveness of public health interventions on SARS-CoV-2 transmission. The Institute for Health Metrics and Evaluation (IHME) [117] used a susceptible-exposed-infectious-recovered (SEIR) model with statistical methods to forecast state-level estimates with current social distancing, face mask use, and testing strategies. Eikenberry et al. [46] used an SEIR model to explore the impact of face mask use by the general public. The COVID-19 Simulator Consortium [39] combined a machine learning method with an SEIR model to consider various features, including geographic and demographic data, and NPIs at the county level.

Agent-based simulation has been employed due to its ability to model detailed interactions of NPIs in populations with heterogeneous demographics and contact networks [3, 81, 86, 92, 126]. Advantages of using an agent-based simulation model include increased flexibility to model

individuals' behaviors and their interactions with environmental constraints such as health system capacities and the ability to incorporate tailored policy recommendations. Early work in the COVID-19 context includes that of Kucharski et al. [81], who simulated the effectiveness of individual level interventions and physical distancing on reducing viral transmission in the United Kingdom. Aleta et al. [3] simulated 2 social distancing relaxation strategies with increased testing, contact tracing, and home quarantine in the Boston metropolitan area.

Finding optimal combinations of NPIs that can control the spread of SARS-CoV-2 is important. Interventions that have been adopted include social distancing measures such as stay-at-home orders, mass gathering bans, work-from-home orders, face mask use, school closures, home quarantine, testing, and contact tracing. However, after easing some restrictions, many regions are experiencing new surges in infections. How to balance the combinations of interventions is challenging for policymakers as they try both to minimize social disruptions and to keep the population safe. More specifically, it is important to study what combination of NPIs should be recommended and when, where, and to what degree interventions are warranted to control the outbreak.

Our work aims to investigate the impact of NPI combinations on reducing SARS-CoV-2 transmission and to identify optimal strategies to control outbreaks in a large urban area. In the United States, Washington State and the City of Seattle have demonstrated some success in controlling the spread of COVID-19 using a 4-phase approach to reopen society devised by the Washington State governor's office in consultation with public health leaders [145]. We develop and calibrate an agent-based simulation model to represent the greater Seattle area, using a synthetic population of King County, Washington. Next, we simulate various levels of NPIs and identify combinations that can contain COVID-19, as well as a phase based plan through the first half of 2021, as vaccinations are not expected to be widely available before that time [105]. This research quantifies the effects of NPI combinations to provide guidance for state and local governments in implementing policy changes to reopen society.

2.2.2 Method

Agent-Based Simulation Model

Our agent-based model of SARS-CoV-2 transmission is a modification of the open-source A Framework for Reconstructing Epidemic Dynamics (FRED) model [56]. The agents in the model represent individuals and are based on population demographics and contact networks. Our

COVID-19 disease model uses established epidemic parameters for disease progression [8, 12, 23, 41, 58, 82, 135, 152]. We incorporate NPIs (e.g., social distancing, face mask use, school closure, testing, and contact tracing) that influence individuals' behaviors and thus affect the viral transmission (Figure 2.1). To represent population demographics and contact networks, our model uses the Synthetic Population Database developed by RTI International [146]. The synthetic population is generated from the 2010 US census data of King County, Washington, and includes approximately 1.9 million individuals with demographic information (e.g., gender, age, and race) as well as contact network information (e.g., location and size of house hold, neighborhood, workplace if the person is employed, and school if the person is a student). For each day in the agent-based simulation, individuals can interact with other individuals in their household, neighborhood, workplace, and school. Specifically, each individual contacts household members with some probability and contacts a random number of members (which we call "contact rates per day") who are in the same neighborhood, workplace, and school.

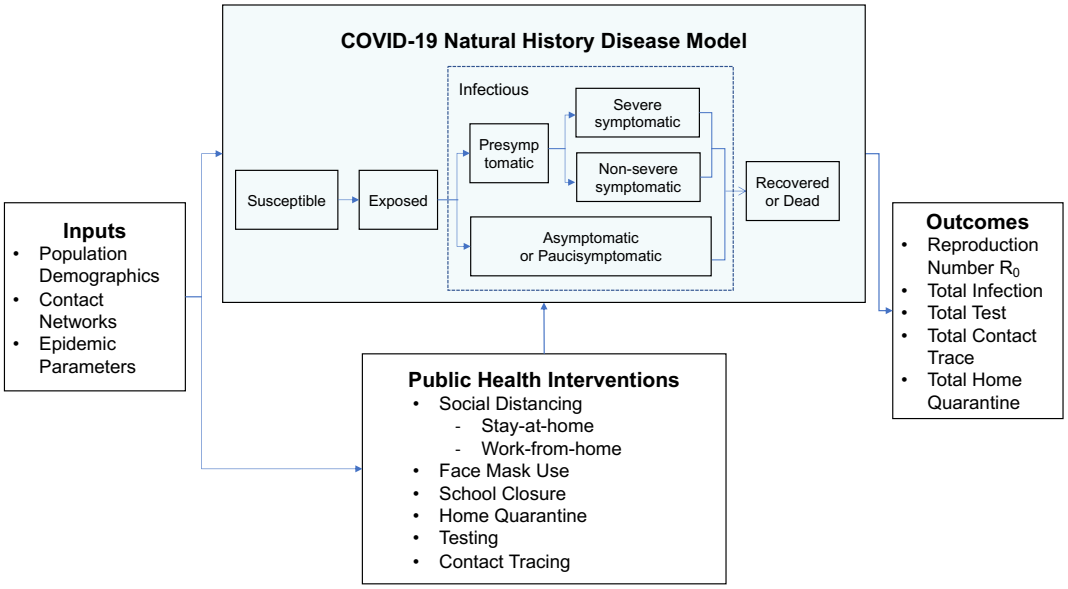


Figure 2.1: Model Overview.

The natural history and transmission of SARS-CoV-2 follows a modified SEIR model. When a susceptible individual comes in contact with an infectious person, there is a transmission probability that the susceptible person becomes exposed to SARS-CoV-2. Once exposed, that person

becomes infectious after a latent period. Depending on their age, the individual becomes presymptomatic or asymptomatic/paucisymptomatic. A symptomatic individual is first in a presymptomatic period, during which the individual is infectious but does not yet show symptoms. Then, the person may have severe symptoms or non severe symptoms depending on their age. Once the infectious period is over, the individual either recovers with assumed full immunity or dies. We did not differentiate between recovered and dead individuals in the model as both types are not driving viral transmission, and we can post calculate mortality. See Appendix A.1.1 for details on disease parameters.

For an infectious individual, the probability that other susceptible members from household, neighborhood, workplace, and/or school become infected depends on the transmissibility of SARS-CoV-2, contact rates at the location, the infectious individual's infectiousness, and the susceptible person's susceptibility to the virus. We modeled three NPIs that reduce contact rates or transmissibility of SARS-CoV-2 and that have frequently been the focus of policy interventions in the United States: social distancing, face mask use, and school closure. In our model, social distancing includes work-from-home and stay-at-home orders, as a result of which a percentage of workplaces is closed and a percentage of individuals does not participate in neighborhood activities. In addition, we considered the effects of face mask use in neighborhoods, workplaces, and schools. If infected individuals wear masks, their infectiousness is reduced, and if susceptible individuals wear masks, their susceptibility to SARS-CoV-2 is reduced. Last, we considered school closure or operating in a hybrid model to decrease transmission among children. School closure becomes more effective if children are the major contributors to community transmission [41]. The objective of these NPIs is to reduce the overall contacts and degree of transmissibility within a community.

We also modeled NPIs related to symptoms or concern about exposure, which include home quarantine, testing, and contact tracing. Without public health messaging about quarantines, some symptomatic individuals may decide themselves to quarantine at home if they have symptoms. Since the tendency to stay home when symptomatic may differ by regions, we calibrated this behavioral parameter for King County (see parameter calibration in Appendix A.1.1). Testing and contact tracing can identify infectious individuals and their close contacts, which may lead to an increase in the home quarantine percentage of infectious and potentially infected populations. If the close contacts are identified by case investigators, each person is recommended to get tested and be home quarantined for two weeks since last exposure [25]. The effectiveness of contact tracing depends on the ability to identify all contacts and the ability of those who are identified and

contacted to comply with testing and home quarantine. In our model, a percentage of symptomatic individuals is tested. If an individual tests positive (i.e., is a potential infector), the person is home quarantined and the person's recent close contacts may be traced and home quarantined, based on compliance parameters. See Appendix A.1.1 for detailed model specification and data input.

The main outcome of the model is number of infections, including daily new infections and total infections. Other outcomes include peak incidence, average daily testing and contact tracing resource usage, and total number of home quarantines.

Calibration Procedure

We calibrated transmissibility of SARS-CoV-2, daily contact rate at each location (household, neighborhood, school, and workplace), and default home quarantine percentage of symptomatic individuals. We used a 2-stage approach to target the epidemic in King County from January 15, 2020, through May 31, 2020. We targeted parameter sets that satisfy the basic reproduction number (R_0) of COVID-19, infection fatality ratio (IFR), and infection fatality ratio of symptomatic individuals (IFR-S) to known literature ranges. IFR is defined as "the number of individuals who die of the disease among all infected individuals (symptomatic and asymptomatic)," and IFR-S is defined as "the number of individuals who die of the disease among all symptomatic individuals" [23]. See detailed description of the calibration procedure in Appendix A.1.1.

Intervention Parameters

We categorized governmental social distancing orders based on the severity of the order, where open, low, medium, and high categories correspond to 0%, 20%, 50%, and 80% of coverage, respectively. For example, when a low social distancing order is enacted, 20% of workplaces remain closed and 20% of the population do not have neighborhood contact. In our scenario analysis, we also considered phase-based social distancing in which the degree of social distancing is decided based on the last 2 weeks' diagnosed cases, and each policy change is maintained for at least 3 weeks. For example, if a community was in phase 2 with medium social distancing for more than 3 weeks and if the number of diagnosed cases per 100,000 remained between 10 and 25 in the following 2 weeks, the community is allowed to progress to phase 3 with low social distancing. Face mask use was modeled with two parameters, compliance and effectiveness. According to Eikenberry et al. [46], if typical masks are worn, they reduce transmissibility by 50%. In our

model, if an infector uses a face mask, the infector's infectiousness is reduced by 50%, and if a susceptible individual uses a mask, the individual's susceptibility is reduced by 50%. Last, we assumed that schools operate in a hybrid model from September with reduced contact rates. See Appendixes A.1.3 and A.1.4 for additional details.

Furthermore, we assumed that all severely symptomatic individuals and some mildly symptomatic individuals would get tested. Asymptomatic individuals are not tested unless contact traced. We further assumed that all confirmed cases proceed to home quarantine and 90% of the cases comply with providing their contacts for the last 2 weeks [145]. Contact tracing effectiveness was defined as the probability that households, schools, workplaces, and neighborhood members comply with home quarantine and testing when contact traced. We assumed the testing and contact tracing capacity is sufficient. Anyone who is home quarantined has only limited household activities (50% reduction in household contact). See Appendixes A.1.3 and A.1.4 for intervention parameters and transmission equations.

Analyses

We simulated the model for 500 days since the first day of infection in King County (January 15, 2020) to May 28, 2021, as widespread coronavirus vaccinations are not expected to be available until late spring of 2021.22 The interventions from January 15, 2020, through May 31, 2020, were based on policies mandated in King County. From June 1, 2020, to May 28, 2021, we simulated with parameter settings in Table 2.1, where the base case used the bolded parameter settings (see Appendix A.1.2 for the timeline of the base case). For the base case and all scenario analyses, we replicated 50 simulation runs with a different random seed. We reported outcomes including median values of infection numbers and plotted 7-day moving-average values of daily infection numbers.

We first conducted 1-way scenario analyses on compliance with NPIs, daily imported infections, and intervention parameters such as testing delay and face mask effectiveness. Next, we conducted scenario analyses on 4 intervention parameters to determine optimal combinations of interventions that can control the disease (Table 2.1). Specifically, we conducted 2-way scenario analyses between 2 NPIs to observe their interactions in driving down total infections. Furthermore, we investigated all combinations of the 4 NPIs in a 4-way analysis of 300 policy scenarios. Last, we considered phase-based social distancing in which social distancing measures are decided based on the current epidemic trend, and we compared results from phase-based social distancing

with low social distancing (the base case) and medium social distancing. When evaluating these scenarios, we compared simulated total infections from June 1, 2020, to May 28, 2021. Specifically, we determined NPI combinations that satisfy Washington State's objective of having fewer than 25 diagnosed cases per 100,000 people per 14 days [145]. We translated this objective into total infections from June 1, 2020, to May 28, 2021 (362 days), which resulted in identifying NPI combinations with total infections under 30,600 cases during the period (see Appendix A.1.3 for the calculation).

	Parameter Settings[a]
Social Distancing	Open (0%), **Low (20%)**, Medium (50%), High (80%)
Face Mask Use	0%, 25%, **50%**, 75%, 100%
School Closure	Open (0%), 25%, **Hybrid (50%)**, 75%, Closed (100%)
Testing and Contact Tracing[b]	Low, **Medium**, High

[a] Base case is in bold.
[b] See Appendix A.1.1 for details.

Table 2.1: Scenario analyses for levels of interventions

Model Validation

To ensure "face validity" of the model, we discussed model results with local public health leaders and external modeling teams at a Washington State Department of Health meeting to review for reasonableness and credibility of the model outputs. We also validated the simulated outbreak trends against estimated real outbreak trends between June and October 2020.

2.2.3 Results

Calibration Result

The calibration process identified two clusters of parameters that both satisfy the calibration target ranges but with different, both plausible, features. The calibrated values for transmissibility of SARS-CoV-2, daily contact rate at each location, and default home quarantine percentage of symptomatic individuals in clusters 2 and 6 are shown in Table A.1.2 in Appendix A.1.1. Cluster 2 represents higher contact rates in household, neighborhood, and workplace. Cluster 6 represents higher overall transmissibility of SARS-CoV-2, school contact rates, and default home quarantine

percentage when symptomatic. Because we could not choose between these 2 clusters of good-fitting parameters, we present results from both and identify similarities/differences in the insights of the analysis. More details can be found in Appendix A.1.1.

Base Case Result

With social distancing relaxed to low (20%) and schools partially reopened in the fall 2020, a second surge is inevitable even if face masks are used and contact tracing is enabled (Figure 2.2). While we calibrated through May 31, 2020, our model's simulated daily new infections show similar trends to the estimated real number of daily new infections until the end of October 2020 in King County. We estimated the real number of daily new infections by dividing King County's reported deaths by the IFR (of 0.65%) and estimated the date of infection to be 21 days prior to the date of death. The 21 days is to account for the median number of days from infection to death [23]. Cluster 6's simulated daily new infections show similar trends to real infections from June to October 2020, while cluster 2 shows slightly higher infection numbers. The parameters of cluster 6 and simulated NPIs illustrate King County's steadily decreasing trend before fall 2020. Although social distancing was relaxed starting in June, the schools remained mostly closed and other NPIs such as face mask use and contact tracing controlled the epidemic during the summer, with approximately 15 new cases per day at the end of August. However, as schools reopen in September, new infections start to spike from October 2020 in cluster 6, with relatively high school contact rates.

On the other hand, since cluster 2 has relatively higher neighborhood, workplace, and household contact rates, adding other NPIs could not offset the effects of relaxing social distancing in the summer; the low level of social distancing results in a second surge of daily new infections. Other outcomes such as peak incidence, average daily testing and contact tracing resource usage, and total number of home quarantines are shown in Tables A.1.5 and A.1.6 in Appendix A.1.4

Without any NPIs, nonsymptomatic (pre- or asymptomatic) transmission accounts for 74% and 75.2% of the entire transmission in clusters 2 and 6, respectively. The proportion increases when NPIs are activated because more symptomatic individuals stay at home after testing or social distancing. With NPIs, 84.3% and 89.1% of infection is caused by nonsymptomatic infectors in clusters 2 and 6, respectively.

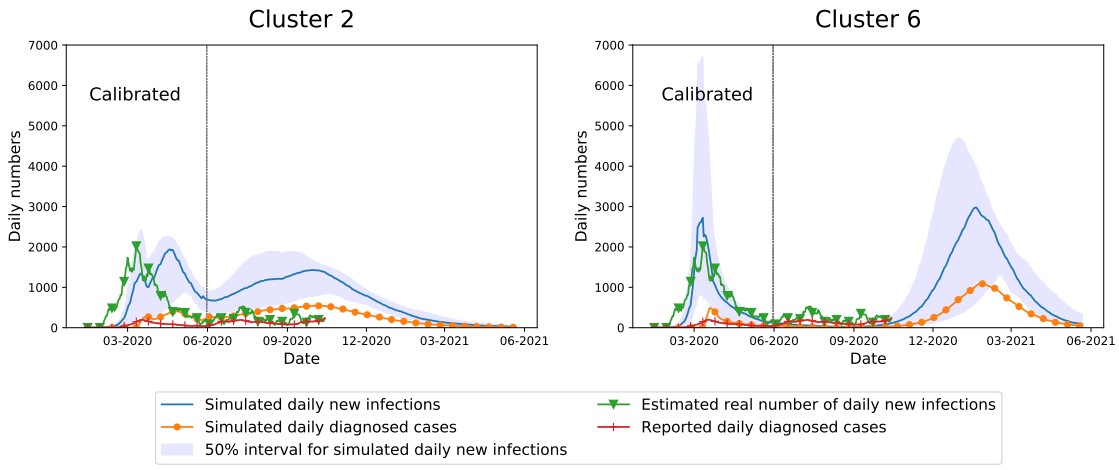


Figure 2.2: Base case result for 2 clusters. The calibration period, from January 15, 2020, through May 31, 2020, resulted in 2 parameters sets, cluster 2 and cluster 6. Our model's simulated daily new infections show similar trends to estimated real number of daily new infections (i.e., reported death/infection fatality ratio [IFR]) until the end of October 2020 in King County.

One-Way Scenario Analysis

Social distancing levels and face mask compliance levels highly affect total infections. In both clusters, medium and high social distancing or 75% and 100% face mask use could achieve Washington State's objective (Figure A.1.7 in Appendix A.1.4). Additional school closure in cluster 6 satisfied Washington State's objective due to its high contact rates in schools, but cluster 2 failed to achieve this objective. In both clusters, testing and contact tracing could not substantially reduce viral transmission. Other one-way scenario analyses, including varying daily imported infections since June 1, 2020, and other intervention parameters (e.g., delays in testing and con tact tracing) yielded similar results to the base case. Face mask effectiveness, on the other hand, had a significant impact on viral transmission, which was similar to changing compliance in face mask use. Since both face mask effectiveness and compliance with face mask use affect transmissibility of SARS-CoV-2, we varied compliance with face mask use in the 2-way scenario analyses and maintained face mask effectiveness at its base case value of 50%.

Two-Way Scenario Analysis

The result of our 2-way scenario analysis between pairs of NPIs are shown in Figure 2.3. As the effects of testing and contact tracing are marginal compared to other NPIs, we only present results between face mask use, social distancing, and school closure. The result between face mask use and social distancing shows that low social distancing and 50% of the population wearing face masks are more effective than medium social distancing and no face mask use (0% compliance) in cluster 6 (Figure 2.3B) and slightly more effective in cluster 2 (Figure 2.3A). High social distancing and 25% face mask use, medium social distancing and 50% face mask use, low social distancing and 75% face mask use, and open social distancing and 100% face mask use all satisfy Washington State's objective in both clusters.

As can be seen from Figure 2.3C,D, the level of social distancing is a stronger driver than the degree of school closure except after a certain threshold; medium social distancing with base case 50% of face mask use success fully controlled the transmission of infection in any degree of school closure. In cluster 2, with open and low social distancing, the number of infections gradually increases as schools become more open. However, cluster 6 satisfied Washington State's objective when social distancing is low and 75% of schools are closed. Therefore, if a community can achieve low contact rates in school through mask use and distancing, safe reopening is likely to be feasible.

Last, 75% or higher compliance of face mask use successfully controlled viral transmission regardless of the degree of school closure (Figure 2.3E,F). In cluster 2, the effect of face mask use is stronger than school closure. In this cluster, even if schools are completely closed, if the face mask use is under 50%, larger outbreaks were observed in the simulation. On the other hand, in cluster 6, school closure was an effective policy if the population used a face mask at 50% or less. Still, except for 50% face mask use and 75% to 100% school closure, none of the combinations met Washington State's objective if less than 75% of the population used face masks. Additional results between testing/contact tracing and other NPIs can be found in Appendix A.1.5.

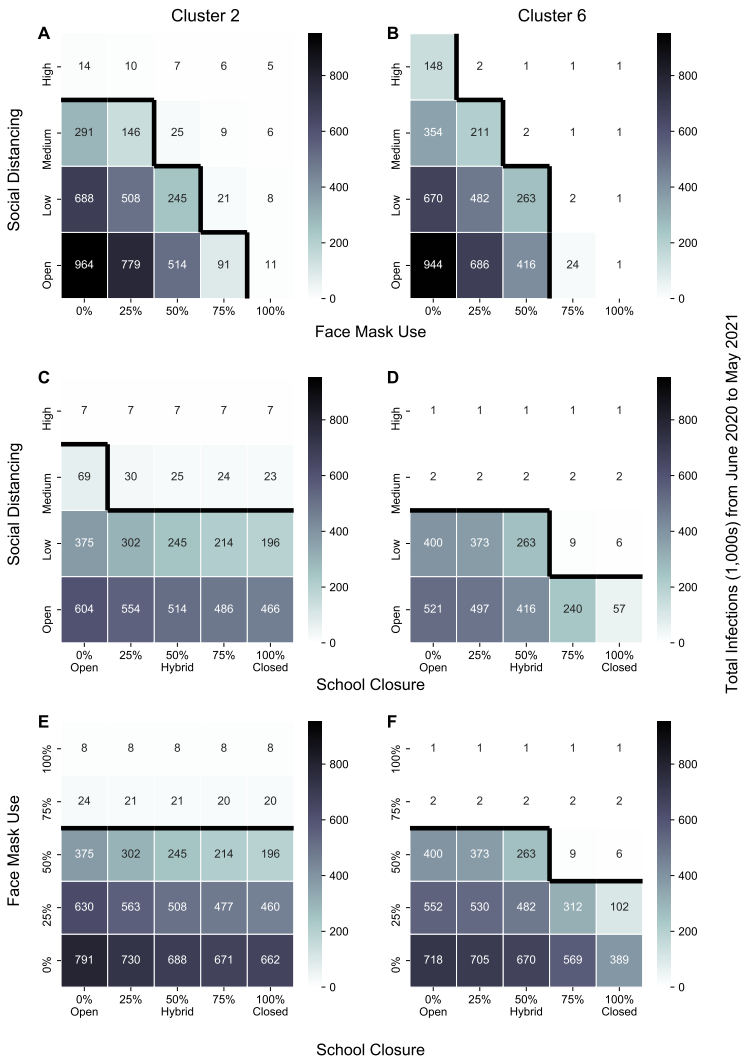


Figure 2.3: Effects of different combinations of social distancing, face mask use, and school closure on total infections (1000s) from June 1, 2020, to May 28, 2021. Black lines indicate combinations that first satisfy Washington State's objective of 30.6 thousand total infections.

Four-Way Scenario Analysis

We simulated all 300 NPI policy scenarios and identified many NPI combinations that reduce total infections to achieve Washington State's objective for total cases. The key message is similar to the 2-way scenario analyses. Overall, high compliance with face mask use is the most important driver keeping total infections low, followed by social distancing level. Society and schools may reopen if 75% to 100% of the population uses face masks. See Appendix A.1.6 for full details.

Phase-Based Social Distancing

We found that phase-based social distancing can effectively eliminate large outbreaks in cluster 2 (Figure A.1.13(A) in Appendix A.1.7), whereas in cluster 6, there is still a second outbreak (Figure A.1.13(C) in Appendix A.1.7). If we maintained medium social distancing from June 1, 2020, to May 28, 2021, then there is no second outbreak in both clusters. In cluster 2, phase-based social distancing resulted in approximately 4 weeks of high social distancing and 3 more weeks of medium social distancing after June 1, 2020 (Figure A.1.13(B) in Appendix A.1.7), which controlled transmission and maintained daily new infections at 48 cases per day during August 2020. With the low prevalence, the phase-based social distancing was able to reduce total infections since June 1, 2020, by 87% compared to the base case. In addition, the phase-based policy had only slightly more infections than medium social distancing only, while neither produced a second wave. In cluster 6, however, although daily new infections remained at 22 cases per day during August 2020 for phase-based social distancing, reopening schools triggered the emergence of a second wave. The increased prevalence required high social distancing until the end of the simulation due to the cluster's higher school contact rates (Figure A.1.13(D) in Appendix A.1.7). On the other hand, maintaining medium social distancing eliminated disease outbreaks before September 2020, and thus no further transmission occurred.

2.2.4 Discussion

Policymakers continue to confront many uncertainties about the best way to control the spread of the novel coronavirus, including when and how to reopen society in the absence of safe, effective vaccines. Our research estimated the impact of NPIs on SARS-CoV-2 transmission and investigated the best policy combinations to control outbreaks in an urban setting. We constructed an agent based simulation model and calibrated behavioral patterns. We found 2 plausible parameter clusters that produced distinct outcome scenarios. In the first scenario, we projected a second wave that begins right after social distancing is relaxed due to high community and workplace contact rates. In the second scenario, the second wave begins with the reopening of schools due to high contact rate in schools. While simulating combinations of NPIs, we explored their interactions and observed threshold combinations that can significantly reduce total infections. In both clusters, medium social distancing and 50% face mask use, or low social distancing and at least 75% face mask use from June 1, 2020, prevented large second waves when schools reopened in hybrid mode

in September 2020. If low social distancing was practiced, at least 75% face mask use may enable reopening of schools with low transmission.

We showed that face mask use played an important role in relaxing social distancing and school closure. Even if social distancing was relaxed from medium to low, total infections remained low if the face mask use was at 75% or higher. In communities with low school interactions, compliance with face mask use is a more effective intervention than school closure. In communities with high school interactions, 25% or more face mask use had more effects than 25% or more school closure. Considering the high societal burden of social distancing and school closure, face mask use is a low-cost and effective intervention. However, we note that implementing effective mask wearing in younger school-age children has practical difficulties, leading to lower effectiveness in real-world settings. Also, while we fixed the face mask effectiveness at 50%, which is considered typical mask wearing14 in the base case analysis, we showed in 1-way scenario analyses that increasing face mask effectiveness could be as effective as increasing face mask use. A recent systematic review and meta-analysis has concluded that "face mask use could result in a large reduction in risk of infection, with stronger associations with N95 or similar respirators compared with disposable surgical masks or similar." [36] However, in some countries where face mask use is high, there was still a winter surge in infection numbers. Therefore, increasing face mask effectiveness, which is influenced by the type of mask and frequency of usage, is very important. We have explored a wide range of face mask effectiveness (25%–100%) in one-way scenario analysis (see Tables A.4.5 and A.4.6 in Appendix A.4).

Consistent with other findings [81, 106], our results showed that testing and contact tracing alone are insufficient to contain COVID-19. Due to high levels of asymptomatic individuals and transmission occurring without symptoms, even the most effective testing and contact tracing strategies quarantined only 40% of infectious individuals in our simulation. However, compared to the least effective testing and contact tracing strategies, the most effective strategies reduced total infections and peak incidence by more than half. Therefore, ongoing efforts are needed to increase the coverage of both testing and contact tracing.

Our analysis of phase-based social distancing policy highlighted the importance of timing in relaxing social distancing measures. If a community interacts primarily at workplaces and neighborhood venues, 4 weeks of high and 3 more weeks of medium social distancing can significantly reduce total infections by 87% and safely reopen schools with low social distancing. This policy produced effects that are similar to maintaining medium social distancing for the whole period.

Another lesson here is the importance of finding the source of transmission in order to determine the most appropriate NPIs. In cluster 6 with higher school interactions, school reopening during a period of low COVID 19 prevalence could still rapidly spread the disease and require more than 6 months of additional high social distancing. Nonetheless, stricter social distancing reduced only 56% of total infections compared to the base case as the intervention only limited neighborhood and work place activities and thus could not reduce transmissions occurring in households and schools. In this case, rather than stricter social distancing, school closure was the more appropriate measure. In partially open schools, eliminating certain types of activities (e.g., singing or sports requiring close proximity) and/or modifying classroom settings could reduce the contact rate (e.g., classes are outdoors or indoors in well-ventilated areas).

Several limitations exist in the model. Although we estimated disease parameters using the best available knowledge, large uncertainties persist about disease parameters such as percentage of asymptomatic infections and transmission occurring without symptoms. Although evidence suggests immunity might decline substantially within a few months [93], recent data suggest that it might last for substantial periods, and reinfections appear to be rare [98]. Thus, declining immunity does not lead to major transmission in our simulation period, which is short term (500 days). Death is postcalculated in our study since the model does not distinguish between recovered and dead. As the medical community learns more about immunity with time, we could enhance our model to separately account for recovered and death. In addition, combining recovered and deaths may undercount transmission since more recovered individuals can block the chain of transmission from an infectious individual, resulting in a lower probability of contacting susceptible individuals, although this impact may be small considering the overall low IFR. The stress on hospital resources and effects of hospital crowding on mortality are not included. We also did not consider virus mutations, which could affect transmissibility and mortality. While we considered wide ranges of R_0 in the calibration, a recent study [74] suggests the value could have been 4.5 early in the pandemic, which is higher than what we used in calibration. In the intervention parameters, we discretized compliance parameters for simplicity and calibrated to the early trend in the epidemic, but in reality, the compliance values may change with the trajectory of the pandemic. Our assumptions about having the same proportion of people in workplaces and people in neighborhood settings complying with social distancing orders, as well as assumptions about everyone who uses face masks using them similarly in neighborhoods, workplaces, and schools, should be reexamined in future studies. We also did not tailor policies to sub populations with different social-economic status and

comorbidity. We did not consider super-spreader events in King County. Last, we assumed that testing and contact tracing capacity are sufficient, but this may not be the case in many regions of the United States.

In summary, this research highlighted the importance of at least 75% face mask use if we want to be able to relax social distancing and reopen schools while controlling SARS-CoV-2 transmission in a large urban area such as greater Seattle prior to large-scale vaccination. We also found that increased testing and contact tracing alone may not be sufficient to control the epidemic without other NPIs. We hope these findings can provide actionable insights on the best NPI combinations to save lives and minimize social disruptions while effective vaccines are rolling out in 2021. This research also highlights the need for continued NPIs as the US population is vaccinated throughout 2021 in light of limited vaccine quantities and anticipated challenges in vaccination scaleup.

2.3 Impact of Nonpharmaceutical Interventions and Vaccination with Viral Variants and Waning Immunity

2.3.1 Introduction

The COVID-19 pandemic continues to ravage populations around the world. Several factors influence this situation, including the emergence of SARS-CoV-2 variants of concern (VOC) such as the Omicron variant that is highly infectious and immune evasive [26], changes in the level of non-pharmaceutical interventions (NPIs), such as reduced mask use and social distancing, and public reluctance to be vaccinated. Understanding the impact of vaccination and NPI policies on COVID-19 incidence and deaths is needed to guide effective interventions.

The immune response to SARS-CoV-2, including VOCs such as Alpha, Delta, and Omicron variants that have resulted in increased viral loads, disease severity, and/or resistance to immunity conferred by previous infection or vaccination [26], is complex. Studies have shown that immunity wanes over time and the degree of waning may differ by vaccine, time since vaccination or infection, variant type, and demographic factors [9, 27, 54, 101, 129]. The immune response to SARS-CoV-2 confers differential protection against infection, severe disease, or death [4, 107, 132, 148]. Protection against re-infection declines faster for the Omicron variant than for the original Wuhan, Alpha, and Delta variants. However, protection against severe disease remains high for all variants, including the Omicron variant [123].

Questions have been raised about the need for periodic COVID-19 vaccinations and which

population groups should receive it. As of May 2024, CDC recommends the updated bivalent COVID-19 booster vaccine to all individuals aged 12 years or older, with at least a 2-month interval since their last dosage. Evidence has shown that it provides better protection against the Omicron variant compared to older vaccines without additional safety concerns [31]. Nevertheless, repeated vaccinations may result in reduced effectiveness and durability of protection [33, 123]. The effects of periodic vaccination should be explored under various immune scenarios, especially given the emergence of SARS-CoV-2 variants that may impact the efficacy.

In this article, our objective is to understand the role of emerging variants, vaccination, and NPI policies on COVID-19 infections and deaths. We aim to identify scenarios in which COVID-19 can be managed such that the death rate from COVID-19 becomes comparable to the combined annual mortality rate from influenza and pneumonia. As a case study for a large urban area, we simulate COVID-19 transmission in King County, WA (greater Seattle) using an agent-based simulation model. Calibrated to local epidemiological data, our study uses detailed synthetic population data and includes interactions between vaccination and specific NPIs while considering waning immunity and emergence of variants. Virus mutation scenarios include twelve combinations of infectivity, disease severity, and immune evasiveness. A highly effective pan-coronavirus vaccine that works against all strains is considered as an optimistic scenario.

2.3.2 Method

Model Overview

Agent-based simulation has been employed in studies [83, 121, 125, 137] to account for heterogeneous individual behaviors and contact networks. Our agent-based model is based on the open-source A Framework for Reconstructing Epidemics Dynamics (FRED) model [56]. As in Lee et al.[86], we modified the FRED model to simulate SARS-CoV-2 transmission in King County, WA with approximately 1.9 million individuals. The NPIs that we model include social distancing, face mask use, school closures, home quarantine, testing and contact tracing as in Lee et al.[86] Our natural history of COVID-19 follows a susceptible-exposed-infected-recovered-susceptible (SEIRS) model framework, including SARS-CoV-2 variants, vaccination, and immunity from natural-infection or vaccination. As shown in Figure 2.4, each disease compartment is stratified by variant type as superscript x, individuals' vaccination history as subscript v, and previous infection history as subscript p. The model is also stratified by age and comorbidities.

These factors affect disease progression, including the probabilities of infection, disease severity, and death.

Values for all parameters and calibration are given in Appendix A.2.1, and detailed transmission equations are given in Appendix A.2.2.

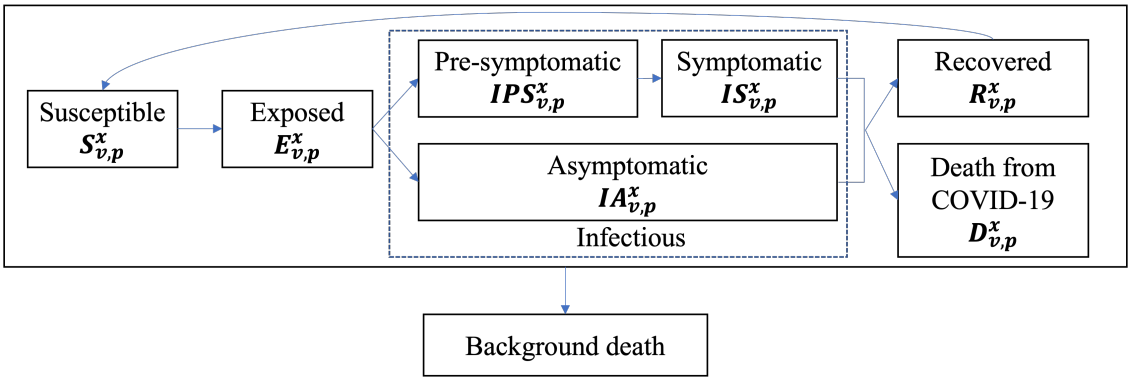


Figure 2.4: Natural history model of COVID-19. In each compartment (S, E, IPS, IS, IA, R, D), the superscript x implies a variant type, and subscripts v and p imply most recent vaccination and previous infection date, respectively.

SARS-CoV-2 Variants We consider SARS-CoV-2 variants that show different infectivity, disease severity, and/or immune evasive properties following previous infection or vaccination. We sequentially introduce three variants that are the most widespread SARS-CoV-2 VOC, specifically, the Alpha, the Delta, and the B.1.1.529 Omicron variants. The variants have evolved to be more infectious while either decreasing or increasing disease severity [44, 49, 51, 67, 133]. The level of natural and vaccine immunity against the Alpha and the Delta variants is assumed to be the same as that of the original Wuhan strain, while the B.1.1.529 Omicron variant has been found to more easily evade the immunity acquired from vaccine or prior infection [64]. We refer to the original Wuhan, the Alpha, and the Delta variant as pre-Omicron strains. In the scenario analyses presented in Section 2.3.2, we introduce potential new variants that mutate every 6 months after the B.1.1.529 Omicron variant, based on data indicating that the previous VOC appearance interval is between 4 and 8 months [142]. See Appendix A.2.1 for details.

Vaccination Vaccination parameters including effectiveness in preventing infection and death are primarily based on the first-generation Pfizer-BioNTech COVID-19 vaccine. In the primary vaccination series (available from January 1, 2021), two doses are administered, with the second dose following the first dose by 21 days. Additional vaccines after the second dose may be ad-

ministered every 6 months. We introduced vaccines to the simulation using age-specific eligibility dates and prioritization policy in Washington State [143, 144]. See Appendix A.2.1 for details.

Immune Response SARS-CoV-2 variants may evade the immune system and increase the probability of infection [97]. We refer to this as 'immune evasion'. Immunity gained from natural infection or vaccination wanes over time and differs by variant types [48, 54, 64]. From clinical studies, we fit a linear regression model to estimate immune evasion [5, 52, 122, 129, 132]. The immunity level has a continuous value ranging from 0 to 1 that depends on individuals' latest infection date, vaccination date, and variant type. In estimating immune evasion, we distinguish whether the immunity was obtained from previous infections or from vaccinations, as research [4, 54, 129] shows that levels of immune evasion vary depending on its source. When an individual has immunity from both previous infection and vaccination, we multiplied their effects based on studies [17, 54, 107] that a hybrid immunity increases protection against reinfection. If an individual is infected, the severity of the disease (probability of dying from disease) may be reduced with previous infection or vaccination. We refer to this as 'immune protection'. We fit a linear regression model to estimate immune protection against death [48, 131, 132]. See Appendix A.2.1 for the equations.

Mortality Deaths from COVID-19 and background mortality are considered. We assume that the infection-fatality ratio of COVID-19 depends on individuals' age and comorbidity status [1, 37]. The ratio can decrease when an individual has immune protection against death from previous infections or vaccinations (Appendix A.2.2). In the agent-based simulation, once an individual dies, the person is removed from each active location (household, neighborhood, school, and/or workplace) and no longer influences future transmission. Background death rate is based on gender and age [50].

Calibration Procedure

We calibrate the model to data for the greater Seattle area from January 15, 2020 to December 31, 2020 by targeting basic reproduction number (R_0) and reported deaths. We fit previous compliance history to NPIs by observing Seattle's sequence of interventions. Parameters that we calibrate include COVID-19 transmissibility, contact rates at each location (household, neighborhood, school, and workplace), and default home quarantine percentage of symptomatic individuals. See Appendix A.2.1 for the calibration procedure and detailed model description.

Parameter Settings for Scenario Analysis

Our simulation period spans four years, from January 15, 2020 (reported first day of infection in King County) to December 31, 2023. The period from January 15, 2020 to December 31, 2020 is used to calibrate parameters. From January 1, 2021 to December 31, 2023 (three years), we perform a scenario analysis by simulating different virus mutations as in Table 2.2 and policy scenarios as in Table 2.3. In all analyses and calibration, we replicate 50 simulation runs. The scenario analysis used common random seeds for variance reduction.

Mutation parameters As listed in Table 2.2, we introduce twelve new variants with mutation scenarios labeled S1 to S12 that vary in infectivity, disease severity, and immune evasion in June 4, 2022 after the B.1.1.529 Omicron variant. We label each mutation scenario as S1 to S12. For ease of explanation, we name the three immune evasion parameter settings as "pessimisti", "neutral", and "optimistic" immune evasion. In the optimistic scenario, we assume that a pan-coronavirus vaccine is always available for all strains, and the vaccine effects always show the same level as against pre-Omicron strains. This provides stronger and more lasting immune protection than our base, first-generation vaccine which is less effective against B.1.1.529 Omicron strains.

Once one of the 12 new variants is introduced to society, we mutate it every 6 months. The mutation parameters follow the same pattern with every 6 months for simplicity. For example, if the virus evolves to increase infectivity by 50% with the same disease severity and immune evasiveness, then 6 months later, the second generation (Dec. 2022 – June 2023) of the new strain will have a 50% further increase in infectivity compared to the first generation (June 2022 – Dec. 2022) of the mutated strain. A third generation (June 2023 – Dec. 2023) mutates once more after 6 months. See Appendix A.2.1 for details.

Vaccination parameters After the 3rd vaccine dose, a periodic vaccination program (4th or higher) is scheduled for every 6 months until the end of simulation. The periodic vaccination program runs from January 23, 2022 to December 31, 2023. When receiving periodic vaccinations, vaccine willingness may decline by 50%, 25%, or remain constant for each additional vaccination in the scenario analyses, as listed in Table 2.3. Individuals' vaccine willingness for primary and booster vaccination depends on their age from King County data [143]. It is assumed that vaccine supply is enough to cover the population, although there may be some delays due to daily limits.

NPI parameters NPIs include social distancing, face mask use, school closure, home quarantine, testing and contact tracing. We simulate different levels of compliance to social distancing,

Mutation Scenarios	Changes in infectivity compared to previous variant	Changes in disease severity compared to previous variant	Changes in immune evasion compared to previous variant
S1	50% more infectious	Same	Pessimistic immune evasion
S2			Neutral immune evasion
S3			Optimistic immune evasion
S4		50% less severe	Pessimistic immune evasion
S5			Neutral immune evasion
S6			Optimistic immune evasion
S7	Same	Same	Pessimistic immune evasion
S8			Neutral immune evasion
S9			Optimistic immune evasion
S10		50% less severe	Pessimistic immune evasion
S11			Neutral immune evasion
S12			Optimistic immune evasion

Table 2.2: Parameter settings for virus mutation scenario analyses. The changes in immune evasion compared to previous variant have three characterizations: Pessimistic immune evasion indicates that the virus mutates to be 25% more immune evasive than is acquired from vaccination or previous infection, Neutral immune evasion indicates that the virus mutates to have the same immune evasiveness as B.1.1.529 Omicron, and Optimistic immune evasion indicates that the virus mutates to have the same immune response as B.1.1.529 Omicron when it is acquired from previous infection but enhanced immune response when it is acquired from vaccination. See Appendix A.2.1 for details.

school closure, and face mask use by aggregating the three factors under one concept—NPI stage (Figure 2.5). As illustrated in Figure 2.5, we consider four levels of NPIs ranging from a fairly closed society (stage 1) to a fully open society (stage 4). We consider three NPI policies as listed in Table 2.3 that consist of different stages and timing as shown in the left panel of Figure 2.5. Timeline 1 gradually opens from 2021 to a fully open society in 2022, whereas Timeline 2 is a moderate NPI policy that opens more slowly and maintains some level of NPIs through 2023. We also introduce a Threshold NPI policy in which the NPI stage is dynamically determined by the last 2 weeks' diagnosed infection cases. Under this policy, the society is in stage 1 if diagnosed cases are higher than 350 cases per 100K population for a 14-day rolling period, in stage 2 if there are 200 to 349 cases per 100K, in stage 3 if there are 100 to 199 cases per 100K, and in stage 4 if there are fewer than 100 cases per 100K. We fix compliance parameters for home quarantine, testing, and contact tracing at the same level as those at the end of 2020.

Combining the nine policy scenarios with the twelve mutation scenarios yields 108 scenarios,

with parameter values given in Tables 2.2 and 2.3.

Policy Scenarios	Parameter Settings
Reduction in vaccine willingness for each additional dose (4th or higher)	50% less severe, 90% less severe
NPI policy	Timeline 1, Timeline 2, Threshold

Table 2.3: Parameter settings for scenario analyses

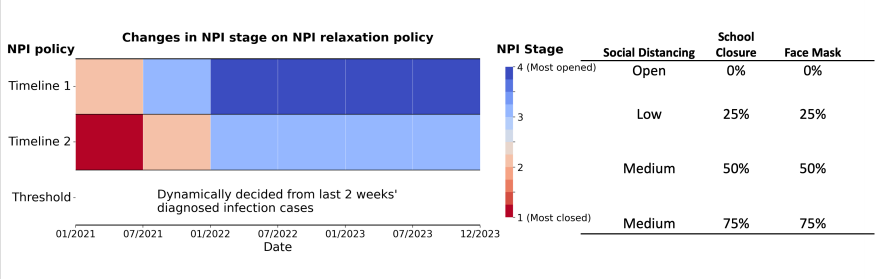


Figure 2.5: Definitions of NPI stages and timeline NPI policies. Date is in month/year.

Output Metrics

We explore the impact of viral mutation, vaccination, and NPI policies on SARS-CoV-2 infections and deaths due to COVID-19. Because new variants after B.1.1.529 Omicron are introduced from June 4, 2022, we focus on deaths that occur from June 4, 2022 to December 31, 2023. Given that the severity of the disease influences public perception and response, we identify scenarios for which COVID-19 deaths can be reduced to or below the mortality rate from influenza and pneumonia, which was 12.6 per 100K population in Washington State in 2017 [28]. This number converts into 374 total deaths from June 4, 2022 to December 31, 2023, or 0.65 person deaths per day in the region.

2.3.3 Result

Increased infectivity and immune evasion are the main drivers of new variants' capacity to dominate other strains.[43, 83] Figure 2.6 presents the number of daily infections by variants under

mutation scenarios S1 to S3 with increased infectivity, and S10 to S12 with reduced disease severity from 2.2. The first row of Figure 2.6 shows that when variants mutate to increase infectivity, the new variants dominate the previous strain. When the variants become more immune evasive (S1, pessimistic immune evasion), the outbreak size is much higher than the outbreak from B.1.1.529 Omicron (Figure 2.6A). If the immune evasion scenario is neutral (S2) or optimistic (S3), the outbreak size is reduced, and its peak is delayed (Figure 2.6B and C). When variants retain the same infectivity with reduced disease severity, dominance of new variants depends on evolution in immune evasion (Figure 2.6D-F). With pessimistic immune evasion (S10), the new variants slowly dominate its previous strain with lower peak size than B.1.1.529 Omicron (Figure 2.6D). However, if variants show the same infectivity and immune evasiveness as B.1.1.529 Omicron (S11, neutral immune evasion) or with better vaccination immune protection (S12, optimistic immune evasion), new variants do not dominate B.1.1.529 Omicron (Figure 2.6E and F).

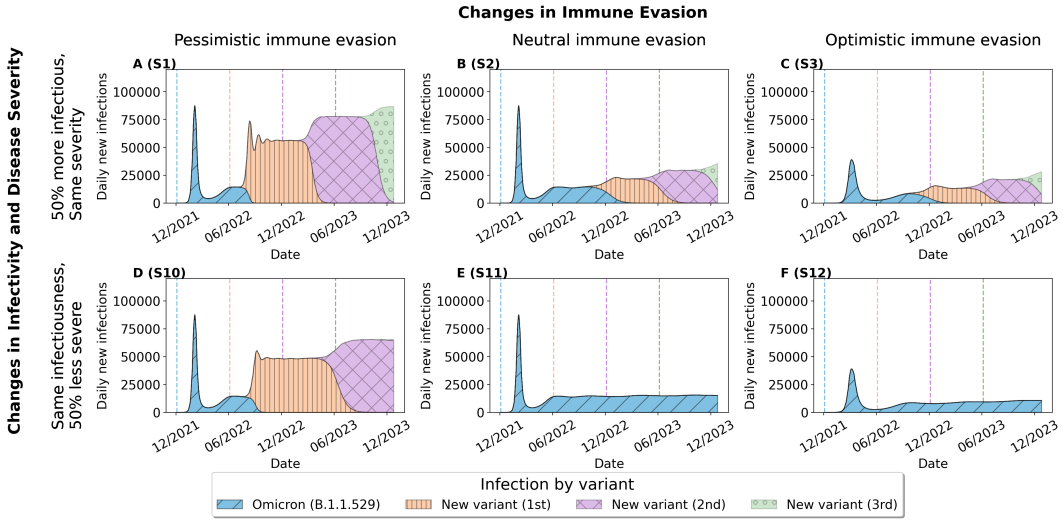


Figure 2.6: Impact of viral mutation on SARS-CoV-2 infections while varying infectivity and immune evasion. Virus mutation scenarios are listed in the upper left corner of each graph. Colored lines in each graph indicate the first imported date of each variant. Periodic vaccination willingness is assumed to reduce by 25% for each additional dose. NPI policy is assumed to be Timeline 1 policy. See Appendix A.2.3 for the results of twelve mutation scenarios from January 15, 2020.

Figure 2.7 presents simulated results for total deaths. The twelve graphs in Figure 2.7 represent mutation scenarios S1 to S12 in 2.2 where the four rows represent changes in infectivity

and disease severity and the three columns represent immune evasion characteristics. Each of the twelve graphs has three markers (squares, circles, and triangles) for vaccine willingness impacting periodic vaccination rate. The three NPI policies are plotted on the horizontal axis, and total deaths (in 1,000s) are plotted on the vertical axis.

Impact of viral mutation on SARS-CoV-2 deaths

In most cases, increased infectivity yields more deaths than mutations that maintain the same infectivity. When the variants' disease severity remains the same and infectivity increases, total mortality increases by 44.7% (Figure 2.7, A vs G), 93.1% (Figure 2.7, B vs H), and 112.7% (Figure 2.7, C vs I) when averaged over vaccine willingness and NPI policies. When the virus mutates to reduce disease severity and the immune evasion scenario is pessimistic, total deaths increase by 16.7% (Figure 2.7, D vs J). On the other hand, when the immune evasion scenario is neutral or optimistic, the total number of deaths decreases by 7.0% (Figure 2.7, E vs K) and 8.3% (Figure 2.7, F vs L), respectively. This is because the new variants with higher infectivity dominate the previous strains, resulting in more infections, but the reduced disease severity leads to fewer deaths.

Similarly, reduced disease severity typically leads to fewer deaths than the same severity. When the virus mutates to increase infectivity, reduced disease severity decreases total mortality by 59.4% (Figure 2.7, A vs D), 52% (Figure 2.7, B vs E), and 56.9% (Figure 2.7, C vs F) when averaged over vaccine willingness and NPI policies. Although the reduced disease severity decreases total deaths by 51.1% (Figure 2.7, G vs J) in the pessimistic immune evasion scenario, the impact of reduced disease severity is negligible when infectivity is the same and immune evasion scenario is neutral or optimistic. The mortality reduces by 0.3% (Figure 2.7, H vs K) in neutral immune case and increases by 0.1% (Figure 2.7, I vs L) in optimistic case. This is because the new variants, which do not mutate to increase infectivity or immune evasiveness, do not replace previous variants as demonstrated in 2.6E and F. Thus, the decreased disease severity of these new variants does not result in a reduction of deaths.

In all instances, increased immune evasion results in a higher death toll. Viruses with a pessimistic immune evasion strategy (first column of Figure 2.7) cause 3,867 average deaths. When the viruses exhibit neutral immune evasion (second column of Figure 2.7), the average total mortality rate decreases by 14.9% from the first column setting. With optimistic immune evasion (third column of Figure 2.7), the death toll further decreases by 30.9% from the second column setting.

Impact of periodic vaccination rate on SARS-CoV-2 deaths

The effect of increasing periodic vaccination coverage heavily depends on new variants' immune evasion property. An effective pan-coronavirus vaccine (S3, S6, S9, S12) and the most optimistic periodic vaccination rate (same vaccination willingness as the booster vaccination rate) yields 1,590 average total deaths, reducing total deaths by 44.6% compared to the most pessimistic periodic vaccination (50% reduction in vaccine willingness for each additional dose). If such a pan-coronavirus vaccine does not exist, and a virus shows neutral immune evasion (second column of Figure 2.7), the average total death reduction is 27.2%. If the virus mutates to evade immunity more easily (first column of Figure 2.7), the death toll reduction is 6.5% when comparing the most optimistic to the most pessimistic vaccination rate. Figure 2.8 illustrates the impact of periodic vaccination on daily deaths when the virus mutation scenario is S1 to S3 with increased infectivity and S10 to S12 with reduced disease severity. The figure illustrates an example in which the benefit of increased periodic vaccination is more apparent when the immune evasion scenario is optimistic.

Impact of NPI policies on SARS-CoV-2 deaths

NPI policies always reduce the death toll regardless of mutation scenarios, as shown in Figure 2.9. Compared to the Timeline 1 NPI policy which fully opens the society from January 2022, the Timeline 2 policy maintains NPI stage 3 from January 2022 to the end of 2023 (see Figure 2.5). With constant moderate NPI policies in Timeline 2, the average total death toll is reduced by 33.6% compared to Timeline 1. The reduction ranges from 21.8% to 47.2% depending on the mutation scenarios. The Threshold policy, which dynamically decides the NPI stage according to the number of cases in a 14-day rolling period, reduces the death toll by 71.3% compared to Timeline 1. The reduction ranges from 55% to 85% depending on mutation scenarios. When the Threshold policy is applied, the NPI stage 1 policy is selected in early 2022 with the rapid spread of B.1.1.529 Omicron variant and remains in NPI stage 1 in most virus mutation scenarios. Even in the case where the virus mutates to be milder, exhibit the same infectivity, and a pan-coronavirus vaccine is available (S12), the society is generally in NPI stage 2, which involves a medium level of social distancing. See Appendix A.2.5.

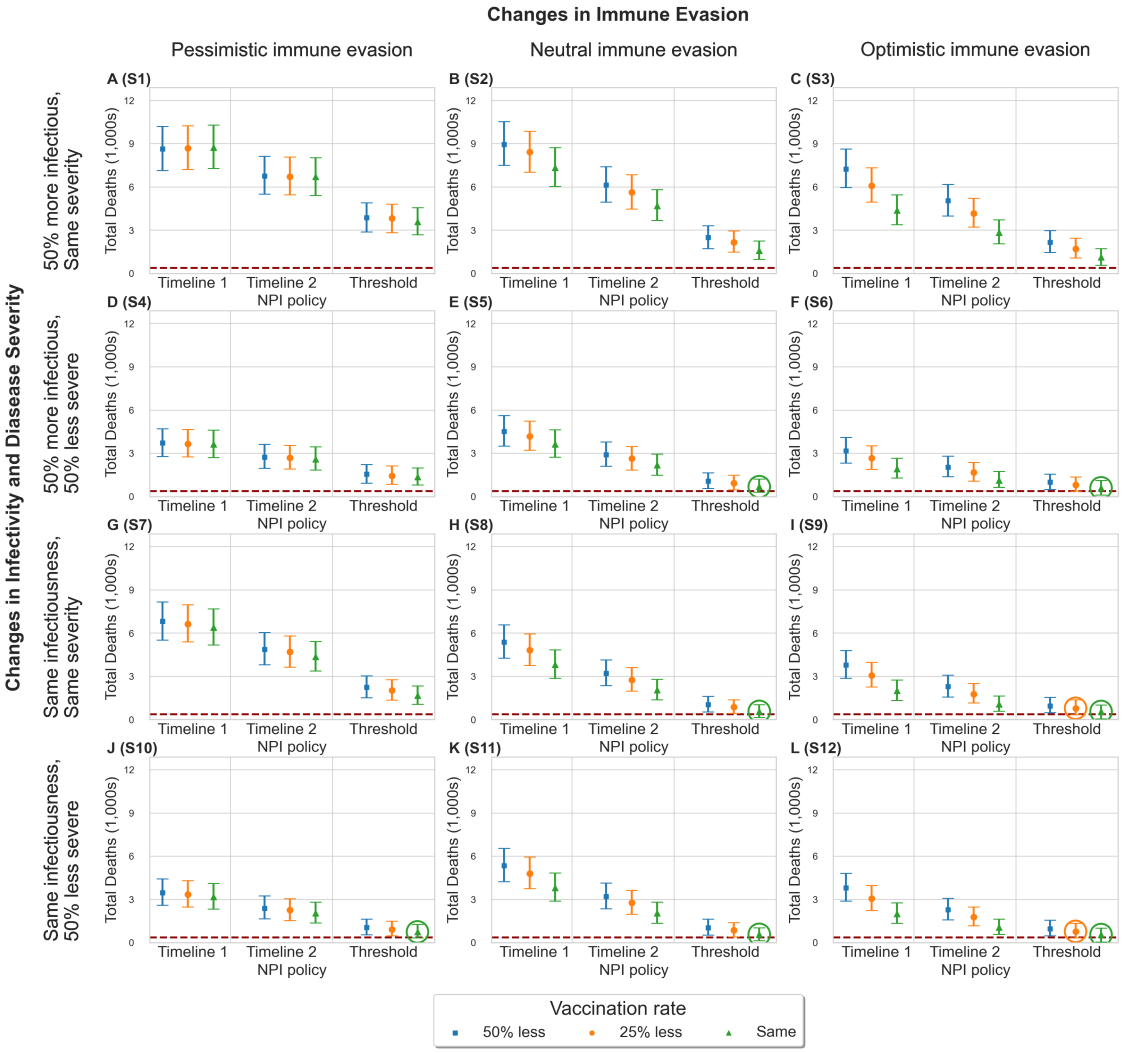


Figure 2.7: Impact of periodic vaccination rate and NPI policies on deaths from June 4, 2022 (first date viral mutation after B.1.1.529 Omicron is imported to the society), to December 31, 2023, with varying mutation scenarios on infectivity, disease severity, and immune evasion. Virus mutation scenario numbers are listed in the upper left corner of each graph. Error bars represent the 25th and 75th percentile values of total deaths, with the dot at the 50th percentile. In each virus mutation graph, red dotted lines represent the number of total deaths from influenza and pneumonia in Washington state in 2017, which is calculated to be 374 deaths during the simulation period. Scenarios with error bars that overlap the objective are circled.

Reducing annual deaths from SARS-CoV-2 to the level of influenza and pneumonia

Our objective is to find scenarios that reduce COVID-19 deaths to the level of influenza and pneumonia, which is 374 total deaths during the simulation period in the region. Out of 108

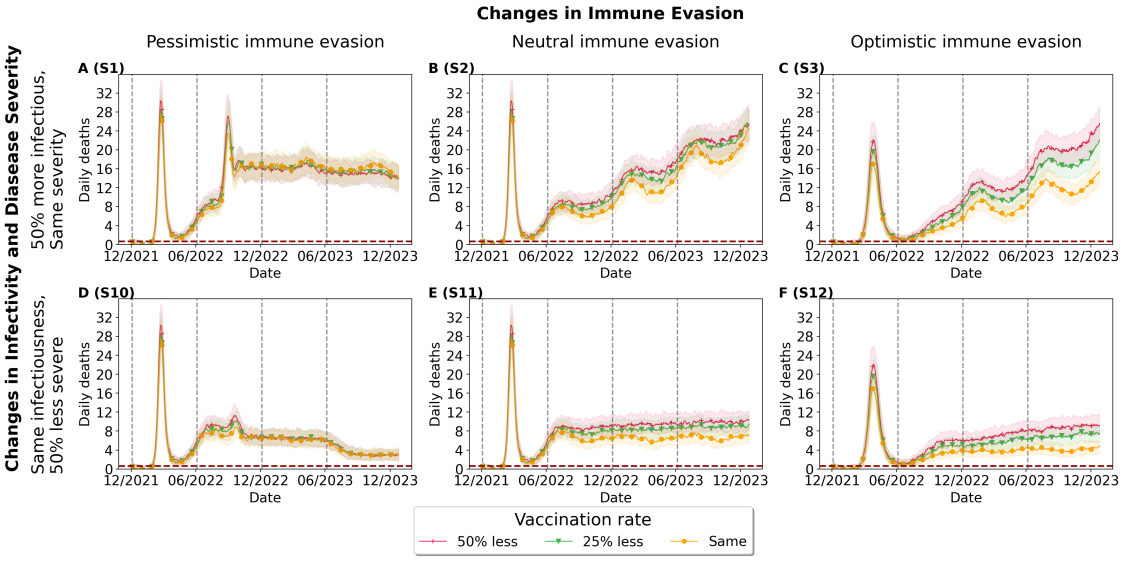


Figure 2.8: The impact of periodic vaccination rate on daily deaths when the virus mutation scenario is S1 to S3 and S10 to S12. The NPI policy is assumed to follow Timeline 1. The shaded area in each line indicates 25th and 75th percentile values of daily deaths. In each graph, the vertical grey line indicates the first imported date of each variant (the B.1.1.529 Omicron, and a new variant that mutates twice more). The red dotted lines represent the number of daily deaths from influenza and pneumonia in Washington state in 2017, which is calculated to be 0.65 deaths per day. See Appendix A.2.4 for the results of all twelve mutation scenarios from January 15, 2020.

scenarios, nine scenarios (with virus mutations S5, S6, and S8 to S12) satisfy the objective within the error bar limits, as indicated by the circled markers in Figure 2.7. All nine scenarios have Threshold NPI policies. Out of the 9 scenarios, 7 scenarios have the same periodic vaccination willingness as the rate of the third booster vaccination. The remaining two scenarios have a 25% reduced periodic vaccination rate (with virus mutations S9 and S12). In Figure 2.9F, in which the virus mutation is S12, the periodic vaccination rate is reduced by 25%, and Threshold NPI policy is applied, the lower error bar also satisfies the objective when converted to daily numbers, equating to 0.65 daily deaths in the simulation period.

2.3.4 Discussion

In this study, we explore the role of hypothetical virus mutation, periodic vaccination, and NPI policies on COVID-19 in a large urban area, King County, WA using an agent-based simulation

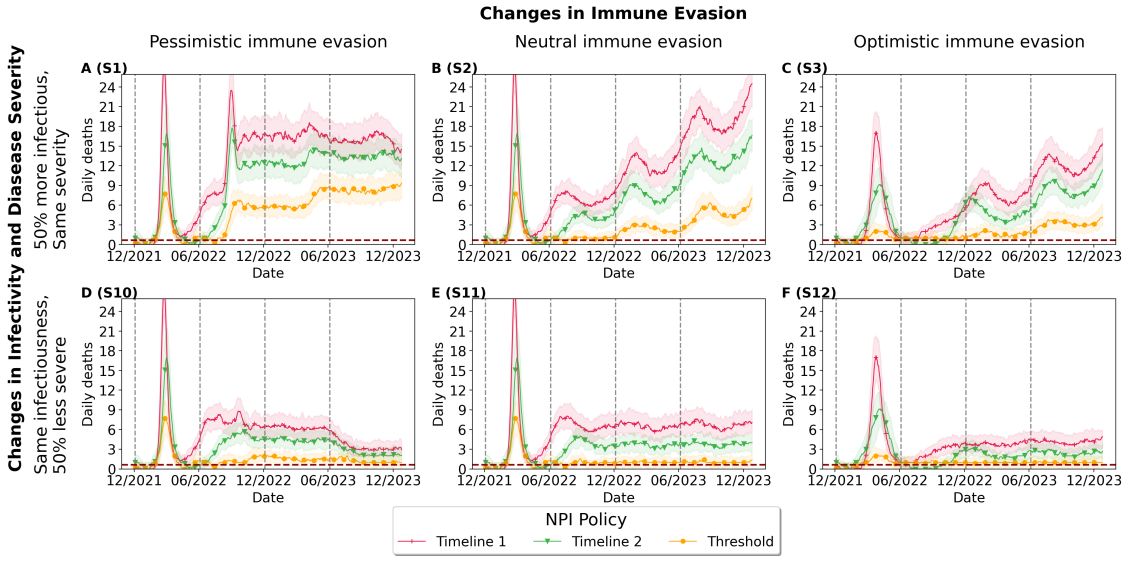


Figure 2.9: The impact of NPI policies on daily deaths when the virus mutation scenario is S1 to S3 and S10 to S12. Periodic vaccination willingness is assumed to reduce by 25% for each additional dose. Shaded area in each line indicates 25th and 75th percentile values of daily deaths. In each graph, the vertical grey line indicates the first imported date of each variant (the B.1.1.529 Omicron, and a new variant that mutates twice more). Red dotted lines represent the number of daily deaths from influenza and pneumonia in Washington state in 2017, which is calculated to be 0.65 deaths per day. See Appendix A.2.4 for the results of twelve mutation scenarios from January 15, 2020.

model. Our study highlights that the effect of increased periodic vaccination coverage on mortality is heavily dependent on the concomitant evolution of immune evasion and is not significant when SARS-CoV-2 mutates to substantially increase immune evasion. In contrast, the effect of strengthening NPI policy is robust to viral mutation. Few scenarios meet the objective of reducing COVID-19 mortality to or below the influenza and pneumonia mortality level by the end of 2023. This raises concerns about managing ongoing COVID-19 community spread using strategies analogous to those for seasonal influenza.

Consistent with other modelling studies [43, 83], our results demonstrate that a novel SARS-CoV-2 strain dominates its previous strain and drives new waves of infections when it has sufficiently increased infectivity or immune evasiveness. Our model indicated that changes in relative infectivity and immune evasiveness determine the dominance of new SARS-CoV-2 strains, irrespective of the mutation scenario. Furthermore, even with changes in vaccination rates and NPIs,

the same dominance was observed, although the time to domination varied.

Our study finds that the majority of deaths comes from people aged 65 years or older, accounting for 67.4% to 81.9% of all deaths in our scenarios. This is in line with CDC's initial decision to recommend the 4th booster vaccine to individuals at high health risk. Furthermore, we observed that increasing the rate of periodic vaccination can reduce the death toll by 27% to 45% as long as the immune evasiveness of new variants remains the same or less than that of the B.1.1.529 Omicron variant. An effective pan-coronavirus vaccine can reduce immune evasiveness and improve protection. The implementation of updated bivalent COVID-19 vaccines, which are believed to provide improved immune response [31, 91], is likely to contribute to reducing the death toll to some extent.

As the COVID-19 pandemic continues, many are fatigued and reluctant to follow restrictive NPI measures, such as social distancing and mask wearing. Now that the COVID-19 pandemic is in its fourth year, our Threshold NPI policy that mostly returns to NPI stage 1, a strict policy, or from late 2022 is likely to be impractical. The criterion in our Threshold policy is based on the number of diagnoses suggested by the WA government before the widespread distribution of vaccination [141]. Now that the disease severity of SARS-CoV-2 is lower, easing the threshold criteria to implement a more relaxed Threshold policy might be a practical alternative.

Agent-based simulation has been employed in studies [83, 121, 125, 137] to account for heterogeneous individual behaviors and contact networks. Our model accounts for real-world evidence of immune response to SARS-CoV-2 so that individuals' level of immune response depends on previous infection and vaccination history as well as variant type and disease outcomes (i.e., infection or death). Moreover, our agent-based model captures individual heterogeneity in behaviors such as mask wearing and compliance with social distancing, and risk factors such as age and comorbidity. Rather than approximating the effects of NPIs as a single variable that changes the force of infection, a method commonly used in mathematical models for simplicity [18, 116, 136], we separately model specific NPIs including social distancing, face mask use, school closure, testing, contact tracing, and home quarantine.

Limitations exist in our model. Some of the virus mutation scenarios may not be biologically feasible. For example, our most pessimistic mutation scenario (S1) that assumes the virus keeps mutating to increase infectivity, increase immune evasiveness, and has the same disease severity may be considered extreme. Our model assumes that the virus follows the same mutation path every 6 months ignoring interactions between other factors. Research shows that high SARS-CoV-

2 incidence rates [127] or infections in immunocompromised individuals [35] could impact the pace and nature of mutation. Recent research has challenged our assumption that hybrid immunity provides stronger protection than immunity generated through either infection or vaccination [30, 139]. Another study suggests that the presence of neutralizing antibodies, i.e., exposure to the current strain of the virus, is a crucial determining factor in the level of immunity [57, 96].

Our model has simplified some individual behaviors. Individuals' compliance with NPIs such as face mask use or social distancing are assumed to be independent of vaccination behavior which might not be true in reality [19, 29]. Future research should consider dynamic human behavior for vaccination and NPIs. Our individuals' vaccine willingness were based on their age and location, but inclusion of other demographic characteristics such as educational status, sex, and political affiliation may be beneficial for a more comprehensive analysis. While some research [104, 120]. highlights sex difference in behavioral responses and clinical characteristics, our study primarily focused on age as a predominant factor influencing human behavior, including vaccination and disease progression [37, 41, 79]. Due to the vast number of possible individual NPI scenarios, we had to aggregate the NPIs to manage a feasible number of scenarios for analysis. We used our previous research [86] to create reasonable aggregated NPIs and timelines.

While we count deaths due to COVID-19 in the simulation, counting excess deaths may give different insights. Our results could overestimate mortality if better antiviral treatments are developed. Seasonal changes in transmissibility or contact patterns might affect the shape of the infection waves. Our calibration was based on the original Wuhan virus and other variants' disease characteristics were based on literature review. We did not calibrate to the current time since our objective was to deliver a broad message on the impact of changes in virus mutation scenarios and vaccination and NPI policies on death toll, rather than predicting the exact outcomes in an urban area.

In summary, our study provides estimated impacts of virus mutation, SARS-CoV-2 vaccination, and NPI policies on COVID-19 outcomes using an agent-based simulation. The development of pan-coronavirus vaccines with increased durability and protection has high potential in reducing the death toll. NPIs are important not only because of their direct effect on reducing COVID-19 infections and deaths but also because of their indirect effect on hindering the emergence of variants by reducing transmission. A dynamic, threshold approach to NPI policy is more effective than fixed policies, implying the need to strengthen surveillance systems for timely reporting of SARS-CoV-2 infections and other communicable diseases with pandemic potential. Few scenarios reduce deaths

from COVID-19 to or below the level of influenza and pneumonia by the end of 2023. However, periodic vaccinations coupled with dynamic NPI policies may succeed in managing COVID-19 as an endemic disease that is similar to seasonal influenza.

2.4 Conclusion

In this chapter, we developed and applied a large-scale agent-based simulation to identify optimal public health interventions that contains the COVID-19 pandemic while capturing heterogeneity in individual health behaviors and viral characteristics. To increase validity, our model is parameterized with epidemiological data and is calibrated to a large urban area (e.g., greater Seattle area) in 2020. In the first study, simulation results highlight the importance of high face mask use compliance and social distancing in the absence of vaccination. Without widespread distribution of vaccination, NPIs should be carefully relaxed to keep infections low. In the second study, we identify optimal vaccination and NPI strategies, considering potential mutations. Our findings highlight the advantages of pan-coronavirus vaccines that provides enhanced and longer-lasting immunity. Regardless of viral mutations, the implementation of threshold NPIs, such as mask mandates and social distancing, adaptively control the COVID-19 mortality risk. Reaching influenza-level mortality remains challenging due to highly contagious variants in our scenario, but combining improved pan-coronavirus booster shots with consistent stricter NPIs shows potential. This research contribute to inform vaccination, NPI, and overall pandemic preparedness decisions. Throughout the chapter, our first objective on designing optimal public health interventions to minimize disease burden of COVID-19 pandemic is satisfied.

Chapter 3

OPTIMIZING VACCINATION CAMPAIGN STRATEGIES CONSIDERING SOCIETAL CHARACTERISTICS

Chapter 2 reinforced the importance of periodic vaccination, especially with vaccines demonstrably effective in preventing infections and reducing disease severity. Yet, vaccine hesitancy continues to be a significant public health challenge. While governments and pharmaceutical companies invest in educational campaigns to promote vaccination, the impact of these campaigns is variable and depends on target population characteristics and messaging strategies. This chapter addresses research objective (2) by identifying optimal vaccination campaigns that may increase vaccination uptake and improve public health.

I develop a networked compartmental model that integrates disease transmission dynamics, evolving vaccination opinions, and targeted campaigns. This approach employs coupled dynamics, evolutionary game theory, and social networks analysis to capture the complex interplay between disease spread and shifting public attitudes toward vaccination. This work advances the modeling of complex individual behavioral responses to vaccination campaigns, expanding upon research that often assumes static vaccination behavior.

Results highlight that campaign effectiveness can range from highly effective to ineffective, or may be unnecessary. This variability depends on societal characteristics like perceived vaccine risks and how individuals balance evidence and social influence in their decision-making. Target demographic groups also adapt to these societal characteristics. This study uses its compartmental model to derive simple yet interpretable optimal campaign strategies, offering actionable insights for real-world practice. The research highlights the importance of understanding societal dynamics to maximize the impact of vaccination campaigns. This research is under review at the Health Care Management Science journal as of May 2024 [84].

This chapter is organized as follows. Section 3.1 provides background on recent vaccination campaign practices and reviews the literature on modeling vaccination behavior within the context of disease epidemics. Section 3.2 provides an overview of the coupled dynamics, optimal vaccination campaign formulation and algorithmic approaches, the numerical experiment setting, and

a sensitivity analysis. Section 3.3 presents the findings of the sensitivity analysis regarding the optimal vaccination campaign allocation and its effectiveness. Finally, Section 3.4 summarizes the findings and discusses limitations and potential improvements for further research.

3.1 Introduction

Although vaccines are effective in preventing infections or reducing disease severity [62], encouraging individuals to get vaccinated is challenging. Individuals hesitate to receive vaccines for various reasons, including effectiveness and safety concerns of vaccination, the inconvenience associated with the vaccination process, and/or the impact of social influencers [88, 118].

In efforts to increase vaccine confidence and willingness, governments often launch public education campaigns to promote vaccination. The impact of these campaigns, however, is not uniform and depends on audience and messaging considerations. According to Steinert et al. [124], randomized controlled survey experiments revealed that the effectiveness of a vaccination campaign varies across eight European countries due to different socio-demographic characteristics and country-level factors. For example, in societies with high levels of misinformation and conspiracy beliefs, messages emphasizing medical benefits may not be effective. In such cases, providing incentives for vaccination or restricting the behavior of non-vaccinated individuals might be more successful. Their findings highlight the importance of tailoring public messaging strategies to specific groups based on their societal characteristics.

The communication channel (e.g., broadcasts, posters, social media influencer, and direct calls) itself can impact the campaign outcomes as well. In the 2021 COVID-19 vaccination rollout, the success of personalized approaches and word-of-mouth strategies in rural U.S. communities illustrates how tailored channels can be impactful in unique populations [14]. In a modeling study, Ho et al. [59] demonstrates that optimizing vaccination incentives based on population characteristics, such as cost sensitivity, can amplify the effectiveness of healthcare insurance incentive programs. These examples emphasize the need for understanding ideal message, channel, and audience for successful vaccination campaigns.

Modeling studies have explored the complex health-related behaviors, such as vaccination decisions, in relation to disease epidemics [13, 42, 68, 72, 95, 151]. These models capture the interplay between human behavior and epidemic, where disease outbreaks influence how people behave, and changes in their behavior, in turn, affect the course of the epidemic. A common approach

to model this dynamic is recognizing that individuals make decisions based on multiple epidemic factors, including perceived disease susceptibility and severity, and the health behavior's perceived benefits and barriers [32]. For example, Durham et al. [42] developed a mask-wearing decision model during the 2003 Hong Kong SARS epidemic, which considered individuals' perceptions on susceptibility and severity to the epidemic, along with the perceived benefits and barriers of mask-wearing.

Social interactions, alongside disease epidemic, influence health behaviors. This influence is modeled by several approaches, often within the framework of social contagion, where individuals imitate behaviors observed through their social connections [13, 68, 69, 151]. This theory assumes that behaviors spread through networks similar to biological transmission. Consequently, established disease transmission frameworks such as the Susceptible-Infectious (SI) or Susceptible-Infectious-Susceptible (SIS) models are applied. For instance, Kandhway et al. [69] applied the SIS and Susceptible-Infectious-Recovered (SIR) framework to model information spread with the goal of maximizing diffusion. Their study tested campaigns like direct recruitment (converting susceptible individuals directly into information-aware) and word-of-mouth incentives (encouraging information-aware individuals to spread the message). Despite optimizing information spread on simplified networks, the approach did not consider individual willingness to accept information.

Some studies integrate social contagion with game theory, where behaviors are contagious, yet individuals have decision-making rationale in their adoption [13, 68, 150, 151]. Bauch [13] introduced an evolutionary game-theoretic model where individuals mimic the decisions of vaccinated individuals through an SI process, and the vaccination behavior change probability is game-theoretic, dependent on perceived risks of illness and vaccine side effects. This approach combines social influence and individual rationality in vaccination decision. Building upon this, Yin et al. [151] refined the model by considering their social neighbors' vaccination decisions within the game-theoretic component. Kabir et al. [68] further expanded the model by considering social influence across individual, neighborhood, and societal levels.

Studies above do not account for how to best distribute resources for optimal vaccination campaigns. Current allocation research focuses on applying and advancing operations research methods to manage resource constraints, but assumes individuals' willingness to vaccinate remains constant [73, 89, 128]. This approach overlooks the importance of dynamic vaccination behavior in designing effective resource allocation strategies. While areas like politics and marketing use models to optimize resource allocation and influence behavior [6, 45, 69], these models often lack

the evolving behavior dynamics and are optimized for simple, randomly generated social networks. This limits their applicability to health policy problems, where individual characteristics such as demographics play a crucial role in outcomes.

Our model integrates the coupled dynamics of disease transmission and vaccination opinion propagation. Given an initial lack of interest in vaccinations, our campaign, similar to word-of-mouth incentives proposed by Kandhway et al. [69], incentivizes vaccinated individuals to promote vaccination within their social networks. This sharing, whether through conversations or social media posts, aims to stimulate consideration of vaccination. While this sharing can trigger thoughts on vaccination, the persuasiveness of the pro-vaccination message that leads to vaccination behavior change depends on multiple factors, such as the current epidemic, vaccination risk perceptions, and vaccination coverage.

Our objective is to identify optimal vaccination campaigns targeting specific geographic regions and age ranges that maximize overall health benefits. While the model is applicable to any infectious disease and vaccination setting, we evaluate and calibrate it specifically for the COVID-19 bivalent booster vaccination campaign in King County, WA in 2023. We discuss societal characteristics that make the model's vaccination campaigns most effective. We also identify optimal demographic targets and emphasize the critical role of tailoring messaging and campaigns to varying societal conditions.

3.2 Methods

3.2.1 Coupled dynamics model on disease transmission and vaccination opinion propagation

The dynamics of disease transmission and vaccination opinion propagation are modeled using a compartmental model, as illustrated in Figure 3.1. The model is based on ordinary differential equations as in Eqs. (3.1).

The model is stratified by four disease compartments, two vaccination opinion compartments, and n demographic groups, resulting in a total of $8n$ compartments. The disease compartments include susceptible (S), infectious (I), recovered (R), and deceased (D) individuals. The vaccination opinion compartments are represented as A, indicating individuals who are anti-vaccination and therefore not vaccinated, and P, representing individuals who are pro-vaccination and have received the vaccination. The n demographic groups combine geographic regions and age ranges. The state variables are represented as the number of susceptible/infectious/recovered/deceased

(S,I,R,D) individuals who hold either anti- or pro-vaccination opinion (A,P) in group i, where $i = 1,\ldots,n$. Table 3.1 lists the state variables and aggregations.

The dynamics of disease transmission follow a Susceptible-Infectious-Recovered-Deceased-Susceptible (SIRDS) model. In group i, susceptible individuals ($S_i^A(t)$ and $S_i^P(t)$) can become infected ($I_i^A(t)$ and $I_i^P(t)$) at the force of infection $\lambda_i^A(t)$ and $\lambda_i^P(t)$, respectively at time t. Recovered individuals ($R_i^A(t)$ and $R_i^P(t)$) can lose their immunity and revert to a susceptible state ($S_i^A(t)$ and $S_i^P(t)$) at immunity loss rates ζ_i^A and ζ_i^P, respectively. Infected individuals ($I_i^A(t)$ and $I_i^P(t)$) may recover ($R_i^A(t)$ and $R_i^P(t)$) at recovery rates γ_i^A and γ_i^P, respectively, or die ($D_i^A(t)$ and $D_i^P(t)$) at mortality rates μ_i^A and μ_i^P, respectively.

Opinion propagation in the model occurs via transitions from anti-vaccination (A) compartments to pro-vaccination (P) compartments. We conceptualize this process as a social contagion, where individuals' vaccination opinions change through social interactions. Similar to infection spread, our model uses an SI mechanism, where "S" represents anti-vaccination individuals and "I" represents pro-vaccination individuals who share opinions. The force of opinion propagation indicates the rate at which anti-vaccination individuals become pro vaccination through opinion sharing. In this context, individuals holding anti-vaccination opinions in group i ($S_i^A(t)$ or $R_i^A(t)$) can become pro vaccination ($S_i^P(t)$ or $R_i^P(t)$) at the force of opinion propagation $\theta_i(x,t)$ at time t given a vaccination campaign x. Here, $x = (x_1,\ldots,x_n) \in \mathbb{Z}^n$ represents a vaccination campaign where x_i is the number of incentivized pro-vaccinators in group i that share their opinion with their neighbors. We use game theory to capture the dynamic force of opinion propagation $\theta_i(x,t)$ in vaccination opinions. Further details in the formulation of $\theta_i(x,t)$ are discussed in Section 3.2.1.

We additionally assume that individuals do not change their vaccination opinion while they are infectious ($I_i^A(t)$). We also assume that the supply for vaccines is sufficient so that individuals who switch from anti- to pro-vaccination immediately receive the vaccine. Once vaccinated, the probability of getting infected, recovering, or dying from the disease is reduced, with respective rates of $\lambda_i^P(t)$, γ_i^P, and μ_i^P. We further assume that individuals who have adopted a pro-vaccination stance do not transition to an anti-vaccination stance within our time horizon.

Details on both disease and opinion progression parameters are explained in their respective sections (Sections 3.2.1 and 3.2.1), with corresponding notation provided in Table 3.2.

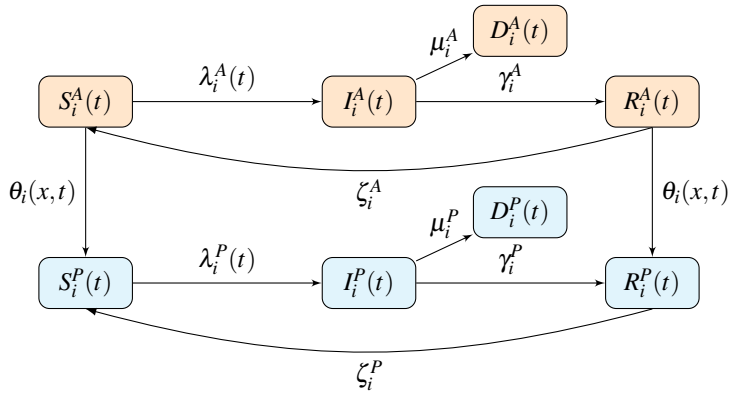


Figure 3.1: Diagram of the coupled dynamics model for group i given vaccination campaign x at time t. Notation is given in Tables 3.1 and 3.2.

The differential equations modeling the coupled dynamics are

$$\begin{aligned}
\frac{dS_i^A(t)}{dt} &= -\lambda_i^A(t)S_i^A(t) - \theta_i(x,t)S_i^A(t) + \zeta_i^A R_i^A(t),\\
\frac{dS_i^P(t)}{dt} &= -\lambda_i^P(t)S_i^P(t) + \theta_i(x,t)S_i^A(t) - \zeta_i^P R_i^P(t),\\
\frac{dI_i^A(t)}{dt} &= \lambda_i^A(t)S_i^A(t) - \gamma_i^A I_i^A(t) - \mu_i^A I_i^A(t),\\
\frac{dI_i^P(t)}{dt} &= \lambda_i^P(t)S_i^P(t) - \gamma_i^P I_i^P(t) - \mu_i^P I_i^P(t),\\
\frac{dR_i^A(t)}{dt} &= \gamma_i^A I_i^A(t) - \theta_i(x,t)R_i^A(t) - \zeta_i^A R_i^A(t),\\
\frac{dR_i^P(t)}{dt} &= \gamma_i^P I_i^P(t) + \theta_i(x,t)R_i^A(t) + \zeta_i^P R_i^P(t),\\
\frac{dD_i^A(t)}{dt} &= \mu_i^A I_i^A(t),\\
\frac{dD_i^P(t)}{dt} &= \mu_i^P I_i^P(t)
\end{aligned} \tag{3.1}$$

where $0 \leq t \leq T$ and $x \in \mathbb{Z}^n$.

Disease transmission parameters

Disease transmission parameters are given in Table 3.2. The force of infection for unvaccinated, anti-vaccination individuals in group i at time t, denoted as $\lambda_i^A(t)$, is the rate at which susceptible

$S_i^A(t)$,$I_i^A(t)$, $R_i^A(t)$, $D_i^A(t)$	Number of susceptible, infectious, recovered, deceased individuals that have anti-vaccination opinion in group i at time t
$S_i^P(t)$, $I_i^P(t)$, $R_i^P(t)$, $D_i^P(t)$	Number of susceptible, infectious, recovered, deceased individuals that have pro-vaccination opinion in group i at time t
$A_i(t)$	Number of anti-vaccination individuals in group i at time t, i.e., $A_i(t) = S_i^A(t) + I_i^A(t) + R_i^A(t)$
$P_i(t)$	Number of pro-vaccination individuals in group i at time t, i.e., $P_i(t) = S_i^P(t) + I_i^P(t) + R_i^P(t)$
$N_i(t)$	Total population in group i at time t, i.e., $N_i(t) = S_i^A(t) + I_i^A(t) + R_i^A(t) + S_i^P(t) + I_i^P(t) + R_i^P(t)$

Table 3.1: Description of state variables and aggregations

individuals become infected. Similar to literature [53, 154], the rate is formulated as

$$\lambda_i^A(t) = \sum_{j=1}^{n} \beta \cdot C_{ij} \cdot \frac{\sum_{k=1}^{n} C_{kj}(I_k^A(t) + I_k^P(t))}{\sum_{k=1}^{n} C_{kj} N_k(t)} \tag{3.2}$$

where β is the transmission rate of infectious individuals (i.e., average number of infected individuals per infectious individual), and C_{ij} is the probability that an individual in group i comes in physical contact with an individual in group j. The force of infection for vaccinated, pro-vaccination individuals in group i at time t, denoted as $\lambda_i^P(t)$, is formulated as

$$\lambda_i^P(t) = v_\beta \lambda_i^A(t)$$

where v_β is the susceptibility reduction factor when vaccinated.

Individuals acquire immunity through natural disease or vaccination. The immunity loss rate is assumed to be the same for vaccinated and unvaccinated individuals. The rate is given as

$$\zeta_i^A = \zeta_i^P = \tau_m^{-1}$$

where τ_m is the average number of days from recovery to become susceptible.

The recovery rate for unvaccinated individuals in group i, denoted as γ_i^A, is assumed to be the

same for vaccinated individuals, and is given as

$$\gamma_i^A = \gamma_i^P = (1 - \alpha r_i) \cdot \tau_r^{-1}$$

where α is the infection-fatality rate of a reference group (e.g., ages 18 to 29 years), r_i is group i's mortality risk compared to the reference group, and τ_r is the average number of days from infection to recovery.

The mortality rate for unvaccinated individuals in group i, denoted as μ_i^A, is given as

$$\mu_i^A = (\alpha r_i) \cdot \tau_d^{-1}$$

where τ_d represents the average number of days from infection to death. The mortality rate for vaccinated individuals in group i, denoted as μ_i^P, is given as

$$\mu_i^P = v_d \mu_i^A$$

where v_d is the mortality reduction factor when vaccinated.

Force of opinion propagation $\theta_i(x,t)$

We model opinion propagation as an SI process, where unvaccinated, anti-vaccination individuals in group i transition to vaccinate, pro-vaccinators at a force of opinion propagation $\theta_i(x,t)$ at time t with a vaccination campaign x. The force of opinion propagation is modeled to parallel the force of infection (Eq. (3.2)). That is,

$$\theta_i(x,t) = \sum_{j=1}^{n} (o \cdot \eta_i(t)) \cdot O_{ij} \cdot \frac{\sum_{k=1}^{n} O_{kj}(\phi P_k(t) + x_k)}{\sum_{k=1}^{n} O_{kj} N_k(t)} \tag{3.3}$$

where o is the average daily vaccination opinion sharing contacts per individual, $\eta_i(t)$ is the probability of accepting pro-vaccination opinions for individuals in group i, and O_{ij} is the probability that an individual in group i comes in communication contact with an individual in group j. We let ϕ denote the proportion of pro-vaccinators who voluntarily share their opinions. Within a vaccination campaign vector $x = (x_1, \ldots, x_n)$, x_k represents the number of incentivized pro-vaccinators in group k that share opinions with their neighbors. Thus, $(\phi P_k(t) + x_k)$ indicates the number of

Disease related parameters	
Disease transmission parameters	
$\lambda_i^A(t)$	Force of infection for disease transmission when unvaccinated in group i at time t
$\lambda_i^P(t)$	Force of infection for disease transmission when vaccinated in group i at time t
ζ_i^A	Immunity loss rate when unvaccinated in group i
ζ_i^P	Immunity loss rate when vaccinated in group i
γ_i^A	Recovery rate when unvaccinated in group i
γ_i^P	Recovery rate when vaccinated in group i
μ_i^A	Mortality rate when unvaccinated in group i
μ_i^P	Mortality rate when vaccinated in group i
Input disease parameters	
C_{ij}	Probability that an individual in group i comes in physical contact with an individual in group j
β	Average number of infected individuals through contact with one infectious person
v_β	Susceptibility reduction factor when vaccinated
v_d	Mortality reduction factor when vaccinated
τ_m	Average days from recovery to become susceptible
τ_r	Average days from infection to recovery
τ_d	Average days from infection to death
α	Infection-fatality rate of a reference group
r_i	Relative mortality risk ratios in group i compared to a reference group
Opinion related parameters	
$\theta_i(x,t)$	Force of opinion propagation in group i at time t with campaign x
Intermediate parameters used in force of opinion propagation $\theta_i(x,t)$	
$\eta_i(t)$	Opinion persuasiveness of an individual in group i accepting pro-vaccination opinion at time t
$J_i^R(t)$	Vaccination probability of group i due to evidence-based factor at time t
$J_i^E(t)$	Vaccination probability of group i due to social influence factor at time t
$B_i(t)$	Perceived benefits of vaccination in group i at time t
Input opinion parameters	
O_{ij}	Probability that an individual in group i comes in communication contact with an individual in group j
o	Average daily vaccination opinion sharing contacts per individual
ϕ	Initial proportion of pro-vaccine individuals who voluntarily share opinions
ρ	Relative weighting on social influence factor in vaccination decision (ranges from purely evidence-based at 0 to purely driven by social influence at 1)
ν	Perceived risk of vaccination
k_R	Sensitivity of evidence-based factor to vaccination probability
k_E	Sensitivity of social influence factor to vaccination probability

Table 3.2: Description of disease and opinion related parameters

pro-vaccinators in group k who share their opinions with their neighbors at time t.

While Eq. (3.3) shares a similar structure with Eq. (3.2), both capturing a contagion process, there exists a distinction. Unlike disease transmission, which is typically modeled with a constant transmission rate β, individuals' decisions to vaccinate involve complex considerations beyond a single transmission rate. To account for this complexity, we introduce $\eta_i(t)$, the probability that an anti-vaccinated individual in group i is persuaded by opinion sharing to change their stance and get vaccinated. We term this concept "opinion persuasiveness," which parallels the notion of susceptibility to infection in disease models.

The opinion persuasiveness $\eta_i(t)$ is modeled by both evidence-based and social influence factors in vaccination behavior. This approach draws inspiration from the evolutionary game theoretic framework presented in Yin et al. [151], and is expressed as

$$\eta_i(t) = (1-\rho)\cdot J_i^R(\nu,t) + \rho \cdot J_i^E(t) \tag{3.4}$$

where ρ is relative weighting on social influence factor in vaccination decisions, $J_i^R(\nu,t)$ is the probability of an individual in group i choosing vaccination due to evidence-based factor at time t, and $J_i^E(t)$ is the probability of an individual in group i choosing vaccination due to social influence factor at time t. Thus, when $\rho = 0$, individuals rely solely on evidence-based factor in their vaccination decision, whereas when $\rho = 1$, they base the decision purely on social influence factor. When $\rho = 0.5$, individuals consider both both evidence-based and social influence factors.

The vaccination probability due to evidence-based factor $J_i^R(\nu,t)$ is modeled by comparing the perceived benefits and risks of vaccination. We model that the primary benefit lies in the reduction of mortality risk from the current epidemic, driven by the vaccine's effectiveness in preventing both infection and deaths. Perceived vaccine risk includes direct risks (the probability of experiencing vaccine side effects) and indirect risks (factors like vaccination costs and potential supply issues). To translate this risk-benefit assessment into a probability between 0 and 1, a logistic regression transition is employed. That is,

$$J_i^R(\nu,t) = \frac{1}{1+\exp\{-(1/k_R)\cdot(B_i(t)-\nu)\}}$$

where k_R is the sensitivity of evidence-based factor, with a higher value indicating a greater responsiveness to changes in evidence-based factor, $B_i(t)$ is the perceived benefit, and ν is the perceived

vaccine risk that we assume is constant in the model. The perceived benefit in [151] focuses on infection risk, whereas our model focuses on the difference in mortality risk between vaccinated and unvaccinated states in the current epidemic, and is formulated as

$$B_i(t) = \frac{\mu_i^A \cdot \left(\lambda_i^A(t) S_i^A(t) + I_i^A(t)\right)}{A_i(t)} - \frac{\mu_i^P \cdot \left(\lambda_i^P(t) S_i^P(t) + I_i^P(t)\right)}{P_i(t)}.$$

The vaccination probability due to social influence factor $J_i^E(t)$ considers social influence, where individuals are more inclined to vaccinate if others within their demographic group do so. Thus, $J_i^E(t)$ is expressed as

$$J_i^E(t) = \frac{1}{1 + \exp\{-(1/k_E) \cdot (P_i(t) - A_i(t)/N_i(t))\}}$$

where k_E is the sensitivity of social influence factor, with a higher value indicating a greater responsiveness to changes in social influence factor. The expression $P_i(t) - A_i(t)$ measures the difference between the pro-vaccination and anti-vaccination population within the group. This difference is then normalized by the group's total population, $N_i(t)$, to determine its relative size.

Using our evidence-based and social influence model, when individuals rely solely on evidence-based factor ($\rho = 0$), vaccination opinion persuasiveness $\eta_i(t)$ increases with higher vaccination benefits, through greater reductions in mortality risk and lower perceived risks of vaccination. With purely social influence factor ($\rho = 1$), persuasiveness increases with rising vaccination rates within their demographic group. In a balanced approach ($\rho = 0.5$), persuasiveness increases with reduced mortality risk, lower perceived risk, and greater vaccination rates within their social group.

3.2.2 Optimal vaccination campaign

With the coupled dynamics of disease transmission and vaccination opinion propagation, we seek an optimal vaccination campaign by incentivizing pro-vaccinators in specific demographic groups to share opinions. With a vaccination campaign x, the opinion sharing population increases, as in Eq. (3.3), which directly strengthens the force of opinion propagation $\theta_i(x,t)$, and thus increases the pro-vaccination population.

Our goal is to maximize the overall health benefits in a community by strategically allocating vaccination campaign resources across different groups. We denote the health objective function as $f(x)$, and quantify this as the increase in the overall pro-vaccination population percentage by

the end of the time horizon. The increase in pro-vaccination population in group i is denoted as $\Delta P_i(x;T)$, and is defined as the difference between the pro-vaccination population of group i with campaign x at time T ($P_i(x;T)$) and the pro-vaccination population at the same time point with no campaign $P_i(\text{nc};T)$. That is,

$$\Delta P_i(x;T) = P_i(x;T) - P_i(\text{nc};T).$$

We denote the total population in group i at time T under campaign x as $N_i(x;T)$. Therefore, our health objective is expressed as

$$f(x) = \frac{\sum_{i=1}^{n} \Delta P_i(x;T)}{\sum_{i=1}^{n} N_i(x;T)}. \tag{3.5}$$

Finally, the optimization problem is formulated as follows:

$$\begin{aligned} &\max_x f(x) && \text{(P)} \\ &\text{s.t. } \sum_{i=1}^{n} x_i = C \\ &\qquad x \in \mathbb{Z}^n \end{aligned}$$

where C is the total number of pro-vaccinators that is incentivized to share their opinion with neighbors, and x is a vector of integers.

Policy identification

The vaccination campaign problem involves nonlinear dynamics with integer decision variables. After experimenting with various black-box optimization approaches, including the trust-region algorithm and genetic algorithm for constrained optimization (as implemented in scipy package [138]), we found that a two-stage coarse grid search was easily explainable and consistently produced the same or better solutions. Therefore, we have opted for this approach in our analysis.

In the two-stage coarse grid search approach, we leverage the fact that our group is stratified by geographic region and age ranges. In the first stage, we optimize allocation for each geographic region and age group independently. The total budget C is discretized into 10% units. In the region-wise approach, we evaluate all possible allocation schemes within the budget, and for

each geographic region, distribute the allocation evenly across age ranges within that region. We then identify the top 10 strategies in the region-wise approach that maximize the problem in Eq. (P). Similarly, all possible age-wise allocations are evaluated, assuming equal allocation across geographic regions, and identify the top 10 age-wise allocation strategies.

In the second stage, we take the cross product of the top 10 allocation schemes for region-wise and the top 10 allocation for age-wise allocation, generating 100 potential allocation strategies. Thus, each demographic group's final allocation is determined by multiplying its relative weightings derived from both geographic region and age factors. To ensure comprehensiveness, we include additional allocations that fully dedicate the campaign budget C to a single demographic group, adding at most n additional potential strategies. All potential allocation schemes are then evaluated to determine the best vaccination campaign allocation.

This approach effectively reduces the problem complexity, which would have a significantly larger number of potential solutions if we had to consider all region and age combinations.

3.2.3 Numerical experiment

Model setting We conduct numerical experiments using data from King County, Washington in the context of the 2023 COVID-19 epidemic with the bivalent booster vaccination. The model runs for one year ($T = 365$) from January 1, 2023, to December 31, 2023 with a daily transition time t. The first two months (January and February) are the calibration period. The remaining months from March to December 2023 are simulated with a vaccination campaign x. The total campaign resources are $C = 20,000$ individuals to incentivize for opinion sharing, which is around 1% of the total population in King County.

The population is segmented into $n = 25$ groups (5 geographic regions and 5 age ranges). King County's ten county subdivisions [108] were grouped into five geographic regions based on both population size and geographic proximity. Ordered by population size, the region groups are Seattle, Seattle East, Federal Way-Auburn-Vashon Island, Issaquah Plateau-Tahoma-Maple Valley, and Enumclaw Plateau-Snoqualmie Valley. The population is also divided into five age ranges: under 18, 18-34, 35-49, 50-64, and 65 and over. For clarity, results are presented using 'region, age' group pairs instead of numeric group i. For example, group (1,3) represents Seattle region and age range 35-49 population.

The interactions among these groups are captured by two parameters: physical connectedness (C_{ij}), governing disease transmission, and opinion sharedness (O_{ij}), governing opinion propaga-

tion. To model physical connectedness, a 25×25 matrix C_{ij}^{Base} was constructed using the Kronecker product of 5×5 regional contact matrix G^{Base} and 5×5 age-wise contact matrix A^{Base}. The regional contact matrix G^{Base} was derived from SafeGraph's mobility phone data [70, 119], and the age-wise contact matrix A^{Base} was based on the United States age-to-age contact patterns sourced from literature [114]. Due to a lack of data on shared opinions across regions and age groups, the base case opinion sharing matrix O_{ij}^{Base} was assumed to generally follow the structure of C_{ij}^{Base}, with the exception that the opinions of children (under 18) do not influence those of older individuals. For more in-depth details on these calculations, refer to Appendix B.1.

Model calibration During the calibration period (January-February), we calibrate four disease related parameters $(\alpha, \beta, \tau_m, s(0))$ and two opinion related parameters (o, ϕ). The parameter $s(0)$ is the initial susceptible population percentage and is assumed to be the same across all groups (i.e., $s(0) = S_i^A(0)/N_i^A(0) = S_i^P(0)/N_i^P(0)$ for all i). The target data comprises eight weekly data on bivalent booster vaccination rate by three age ranges [24], infectious percentage [11, 65, 140], and cumulative deaths [140]. We employed Latin Hypercube Sampling [102] to explore 5,000 parameter sets and identify parameter sets that minimize the mean absolute errors between model outcomes and observed data. Figure 3.2 shows the calibration results. As another check, the 6.7% increase in the simulated vaccination uptake at the end of October without any campaign is similar to observed data of 6.1% increase [24].

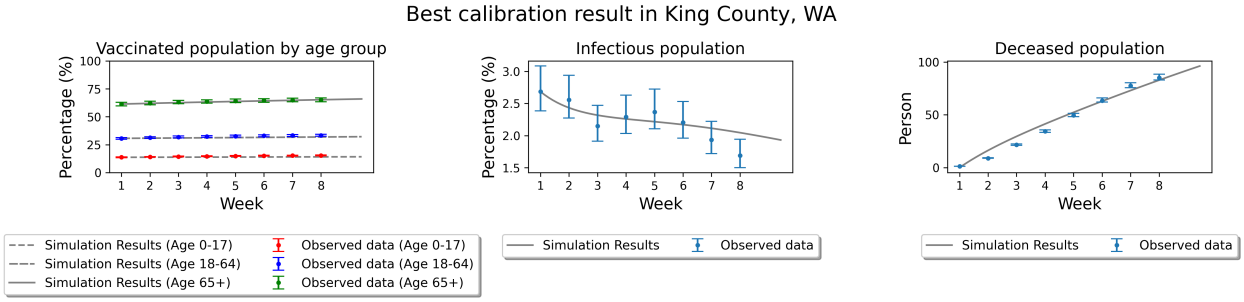


Figure 3.2: Calibration results.

Model execution The model was solved using the deterministic ordinary differential equation solver odeint in the Python scipy package [138]. The data and model used in this research are available on GitHub (`https://github.com/serinlee/Vaccination-Campaign`). For further details regarding the model settings and calibration, please refer to Appendix B.

3.2.4 Sensitivity Analysis

Our model includes several parameters that are meaningful but were estimated or calibrated due to limited real-world data. Thus, we conduct a sensitivity analysis on these parameters, as outlined in Table 3.3.

A crucial parameter influencing vaccination behavior is the force of opinion propagation, denoted as $\theta_i(x,t)$. This rate, as shown in Eq. (3.3), is directly affected by opinion persuasiveness $\eta_i(t)$ which in turn depends on four key societal vaccination opinion parameters: perceived vaccine risk ν, weighting on social influence factor ρ, sensitivity to evidence-based factor k_R, and sensitivity to social influence factor k_E. We place particular emphasis on the key roles of ν and ρ in opinion formation. To explore their combined impact, we conduct a two-way sensitivity analysis for these parameters. Separate one-way sensitivity analyses are performed for the remaining parameters k_R and k_E.

Two-way sensitivity analysis of ν and ρ

We perform a two-way sensitivity analysis of ν and ρ across the ranges in Table 3.3 while setting other parameters at their base values. We ensure realistic experiments by assuming that the upper limit of the perceived risk of vaccination (ν) is set below the initial mortality risk for vaccinated individuals in all groups (i.e., $\nu \leq \max_{i\in[1,n]} \left[\mu_i^P \cdot \left(\lambda_i^P(0) S_i^P(0) + I_i^P(0)\right)/P_i(0)\right]$).

Parameter	**Base**	**Sensitivity Analysis**	
		Lower	**Upper**
ν	0.00015	0	0.0003
ρ	0.5	0	1
C	20,000	10,000	50,000
k_R	$1/20000$	$1/25000$	$1/15000$
k_E	$1/12$	$1/15$	$1/9$
O_{ij}	O_{ij}^{Base}	O_{ij}^{Online}	
Objective function to maximize in Eq. (P)	Vaccination rate increase (Eq. (3.5))	Averted deaths (Eq. (3.6)) Minimum vaccination rate increase across groups (Eq. (3.7))	

Table 3.3: Parameters influencing vaccination opinion dynamics, including base values and ranges for sensitivity analysis.

One-way sensitivity analysis

In the one-way sensitivity analysis, we vary the parameter values specified in Table 3.3, including the campaign budget (C), vaccination judgment sensitivity parameters (k_R, k_E), the opinion sharing matrix (O_{ij}), and the objective function to maximize in Eq. (P).

Since the base case value of $C = 20,000$ was chosen as approximately 1% of the population of King County, we examine $C = 10,000$ and $C = 50,000$ in the one-way sensitivity analysis. The lower and upper values of k_R ($1/25000$ to $1/15000$) and k_E ($1/9$ to $1/15$) are adjusted to ensure plausible 5-10% vaccination increases under a "no campaign" scenario at the end of 2023.

Considering that the campaign can target online social media, we replace the base case opinion sharing matrix O_{ij}^{Base} with a new matrix O_{ij}^{Online}. This new matrix incorporates online social network data, including the Facebook Social Connectedness Index [10] for regional connectivity G^{Online} and social media usage statistics by age range [20, 112] to construct an age-wise contact matrix A^{Online}. The alternative opinion sharing matrix reveals a stronger opinion influence in Region 1 and among age ranges 18-49, but a weaker influence in Region 2 and among age ranges 0-17 and 65+ compared to the original opinion sharing matrix. See Appendix B.1 for details.

In addition to the original objective function of the overall vaccination rate increase, Eq. (3.5), we introduce two additional objective functions. The objective function of overall averted deaths is expressed as

$$\begin{aligned} &\sum_{i=1}^{n} \Delta D_i(x;T) \\ &\text{where} \quad \Delta D_i(x;T) = D_i(x;T) - D_i(\text{nc};T). \end{aligned} \tag{3.6}$$

Another objective function is the minimum increase in vaccination uptake rate across all groups to ensure equitable outcomes. This objective function is formulated as

$$\min_{i\in[1,n]} \{\frac{\Delta P_i(x;T)}{N_i(x;T)}\}. \tag{3.7}$$

3.3 Result

3.3.1 Two-way sensitivity analysis of ν and ρ

Optimal vaccination campaign effects

Societal characteristics like perceived risk ν and the weighting on social influence factor ρ impact vaccination uptake increase, a measure of campaign effectiveness defined by Eq. (3.5). Figure 3.3 shows how these characteristics affect the changes in the vaccinated population over time, both with and without an optimal campaign. The shaded region in Figure 3.3 visually represents the difference in vaccination rates between scenarios with and without the campaign. The difference between these quantities at the end of the simulation's time horizon directly indicates the optimal campaign's effectiveness Eq. (3.5).

The campaign is most effective when society has moderate levels of perceived vaccine risk and balanced considerations on both evidence-based and social influence in their decision-making (Figure 3.3e). In this case, the campaign leads to a 19.7% absolute increase in the total vaccinated population percentage by the end of the time horizon. When individuals base decisions solely on social influence factor (Figures 3.3c, 3.3f, 3.3i), the optimal campaign boosts vaccination rates by 10.6% in societies. Since these decisions depend entirely on group vaccination rates, perceived vaccine risk becomes irrelevant, explaining the identical results across the values of ρ.

There are two circumstances where the campaign has limited impact. The first case is when the vaccination uptake is high without any campaign effects, in which case a campaign is unnecessary. In our model, this occurs when the perceived risk is low and evidence-based factor is considered (Figures 3.3a, 3.3b). In this case, the voluntary opinion sharing activity from ϕ percentage of pro-vaccinated individuals in persuasiveness is sufficient to convince most anti-vaccinated individuals to vaccinate, limiting the campaign's potential for additional impact.

The second case when a vaccination campaign is ineffective occurs when the perceived vaccine risk is high and individuals only consider evidence-based factor (Figure 3.3g). In this case, individuals are less likely to be persuaded with vaccination opinions due to high concerns regarding vaccination. Since the campaign increases the opinion sharing population by giving incentives, if the persuasiveness is low, simply promoting vaccination has limited impact on changing behavior. This highlights the need to reduce perceived vaccine risk through messaging.

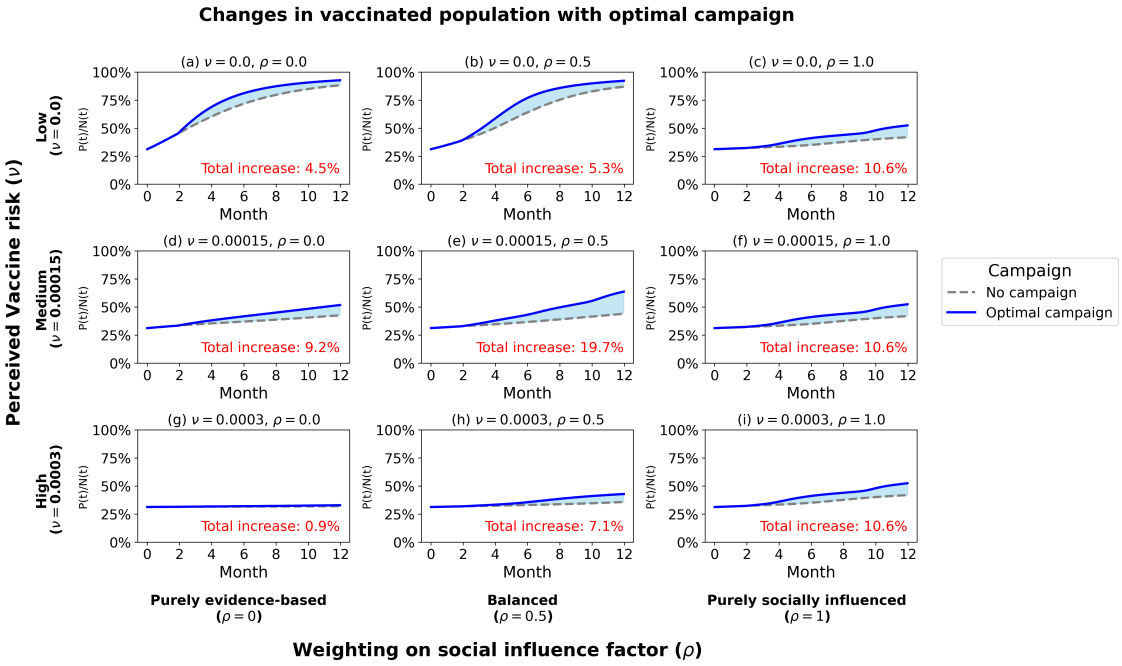


Figure 3.3: Vaccinated population percentage without any campaign and with optimal campaign while varying perceived vaccine risk (ν) and social influence importance (ρ), as in Table 3.3. Dashed lines represent values with no campaign, solid lines represent the optimal campaign results, and the shaded area between them visually represents the changes in values.

Optimal vaccination campaign allocation

As described in Section 3.2.2, our resource allocation strategy combines region-wise and age-wise factors, using a cross-product approach to determine the percentage of campaign resources dedicated to each region and age range. Figure 3.4 illustrates the allocation outcomes using this structure, with the same parameter settings in ν and ρ as in Figure 3.3. For example, in Figure 3.4a, when $\nu = 0$ and $\rho = 0$, the group (Region 1, Age 0-17) receives 25% of the total resources ($50\% \times 50\% = 25\%$), while the group (Region 3, Age 35-49) receives 0% ($10\% \times 0\% = 0\%$). With $C = 20,000$ total resources, the group (Region 1, Age 0-17) gains 5,000 opinion sharers, whereas the group (Region 3, Age 35-49) gains none.

As the perceived vaccine risk increases, the campaign's optimal targeting shifts from 50% of Age 0-17 and 50% of Age 18-34 to 100% of Age 35-49. This allocation leverages the higher opinion persuasiveness of the individuals aged 35-49, stemming from their increased mortality risk and vaccination rate, compared to individuals under age 35. Additionally, this age range have

strong opinion sharing connections with those aged 35 and over (see Figure B.1.2 in Appendix B), amplifying their influence on older demographics. While individuals under age 35 have strong opinion sharing networks, these connections are largely concentrated within their own age range, limiting their ability to influence older individuals. Due to the increasing difficulty of persuading those under 35 as perceived risk rises, campaign resources focus on the 35-49 age range where the potential for vaccination behavior change is greater. Although the 50-64 and 65+ age ranges do not receive direct campaign allocations due to their limited opinion sharing connections, focusing on the 35-49 age range indirectly promotes vaccination uptake within these older age ranges through social connections.

Overall, Region 1, with the highest population density, naturally receives the largest share of resources. However, when vaccination opinions are highly persuasive, such as in Figures 3.4a, 3.4b, resources are shifted towards lower-density Regions 2 and 3. In such contexts, Region 1's high social connectivity leads to faster vaccination uptake without needing all the resources. Meanwhile, less-populated regions need additional support to boost their vaccination rates through increased opinion sharing activities. Even though these regions might have lower influence on opinion spreading, directly allocating resources to these regions ultimately leads to the greatest overall increase in vaccination coverage.

Similarly, distributing resources between Regions 1 and 2 is more advantageous over focusing on Region 1 in some scenarios (Figures 3.4c, 3.4f, 3.4h, 3.4i). With Region 2 representing a 28% of the total population in King County, compared to Region 1's 51%, directing some resources to Region 2 facilitates a notable increase in vaccination uptake within that region. This strategic approach effectively balances the needs of both regions, leading to a greater overall improvement in vaccination rates rather than focusing exclusively on Region 1.

3.3.2 One-way sensitivity analysis

Figure 3.5 displays the results of one-way sensitivity analysis on parameters C, k_R, k_E, O_{ij} and objective function, as listed in Table 3.3. The left panel shows the absolute percent change in vaccination rates compared to a base case (see Table 3.3). This base case is illustrated in Figure 3.3e and Figure 3.4e, where baseline campaign effectiveness is 19.7% and 100% of campaign resources are allocated to group (Region 1, Age 18-34). The right panel shows the optimal allocation by geographic region and age ranges with varying parameters.

Impact of campaign budget (C) Reducing the campaign budget from $20,000$ to $10,000$ re-

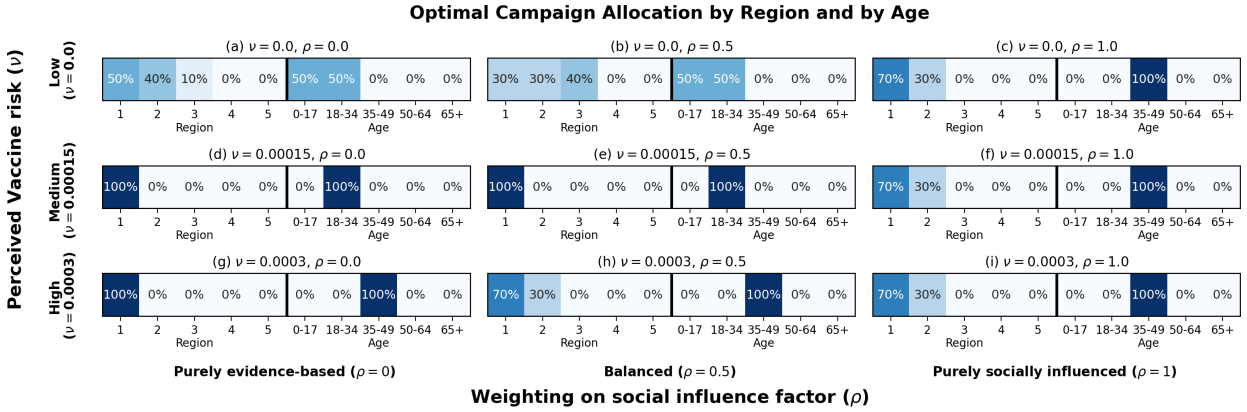


Figure 3.4: The optimal campaign resource allocation by region and age ranges, respectively, while varying societal vaccination opinion characteristics outlined in Table 3.3. The final allocation for each group is calculated by multiplying its region distribution percentage by its age distribution percentage. For example, in (a), the group (Region 1, Age 0-17) receives 25% of the total resources ($50\% \times 50\% = 25\%$), while the group (Region 3, Age 35-49) receives 0% ($10\% \times 0\% = 0\%$).

duces the campaign effectiveness by 11% while the original allocation strategy is maintained. Increasing the budget to $50,000$ boosts vaccination uptake by 12% more than the base case. With greater resources available, 40% of the budget is allocated to Region 2 in regional distribution, and 10% is allocated to the age group of 0-17 in age-wise distribution. This allocation strategy aligns with the approach in Figures 3.4a and 3.4b, where allocating resources to less effective groups maximizes overall uptake.

Impact of sensitivity to evidence-based factor (k_R) and social influence factor (k_E) When society is more sensitive to changes in evidence-based factor ($k_R = 1/15000$), the population is more easily persuaded towards vaccination as the vaccine offers high benefits, owing to high vaccine effectiveness in our setting. In this scenario, vaccination campaigns can boost the vaccination rate by an additional 3% compared to the base case. When sensitivity to evidence-based factor is low ($k_R = 1/25000$), overall opinion shifts more slowly, particularly among those aged under 35. This leads to a slower vaccination uptake, with an 8% decrease compared to the base case. In such scenario, resources are reallocated to prioritize older age ranges, who are more likely to be persuaded and get vaccinated.

Similarly, with high sensitivity to social influence factor ($k_E = 1/9$), vaccination uptake naturally increases at a faster rate even without a campaign. An optimal campaign, with 40% of

resources allocated to Region 2, benefits from this faster vaccination uptake trend, leading to an additional 2% campaign effectiveness. Low sensitivity to social influence factor ($k_E = 1/15$) results in slower overall vaccination uptake change. With the the same resource allocation as in the base case, the campaign effectiveness decreases by 4% from the base case.

Impact of changing opinion sharing matrix O_{ij} Leveraging online social network data for the opinion sharing matrix O_{ij}^{Online} reveals stronger opinion sharing power of group (Region 1, Age 18-49) compared to the baseline O_{ij}^{Base}. While both matrices yield the same budget allocation, using the online sharing matrix results in a 2% increase in campaign effectiveness, due to a greater opinion influence.

Impact of changing objective functions When the objective shifts towards maximizing averted deaths, the allocation shifts towards the 35-49 age range, as they actively spread opinion to older individuals (50 and over). Despite a decrease in overall vaccination uptake (2%), this allocation change results in slightly more lives saved (0.7 person deaths averted).

When the objective is to maximize the minimum vaccination rate increase among groups, the total vaccinated population decreases by 16% compared to the base case. Although overall campaign effectiveness decreases, this approach ensures a more equitable distribution of campaign benefits between groups. The base case allocation results in uneven health benefits. Group (Region 1, Age 18-34) gains the biggest vaccination uptake increase (52.4%), while groups (Regions 4 and 5, Age 0-17) experience minimal gains (0.5% increase). To address this disparity, resources are distributed more broadly across geographic regions and age ranges using the objective function Eq. (3.7). This redistribution leads to more equitable campaign benefits, with a largest vaccination uptake increase of 25.7% in group (Region 2, Age 50-64) and a smallest vaccination uptake increase of 1.4% in groups (Region 4 and 5, Age 18-34).

3.4 Conclusion

Our research explores the dynamic interplay between disease transmission, opinion dynamics, and the effectiveness of vaccination campaigns. By accounting for dynamic vaccination behavior, we aim to develop more effective and targeted strategies for allocating vaccination resources. We find that understanding societal characteristics like perceived vaccine risk and the public's decision-making approach (evidence-based or social influence) is essential for successful campaigns. Results predict that the model's campaigns can influence vaccination decisions but the

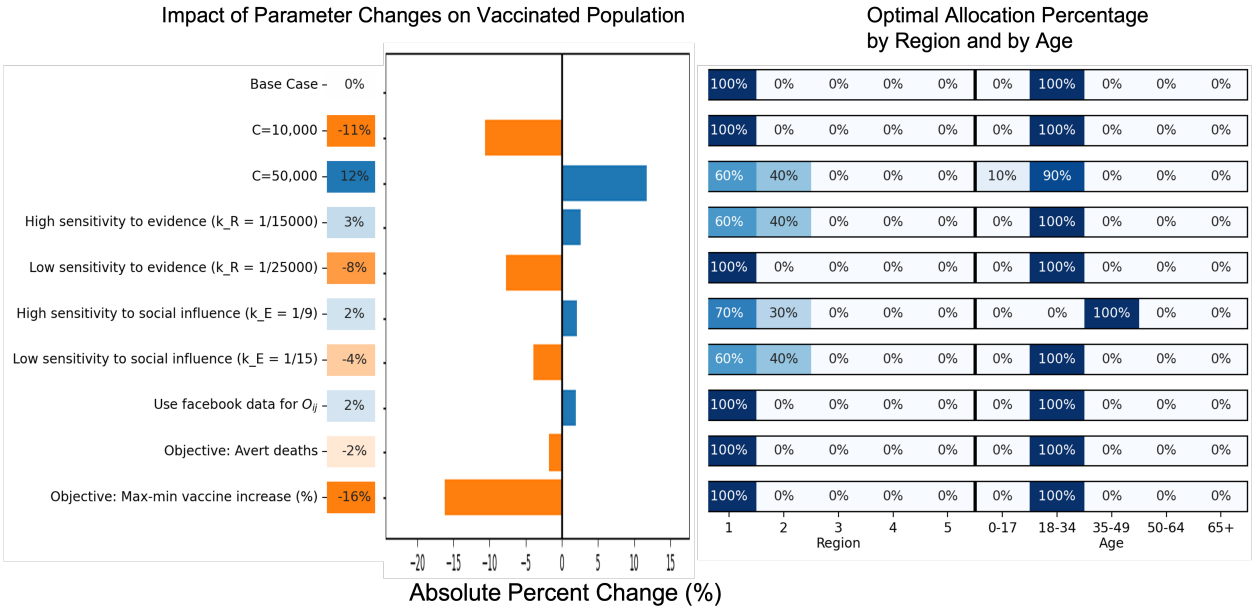


Figure 3.5: One-way sensitivity analysis results with parameters C, k_R, k_E, O_{ij}, and the objective function, as shown in Table 3.3. The left panel shows the absolute percent changes in vaccination population uptake and the right panel illustrates the optimal allocation by geographic region and age ranges when varying the parameters.

campaign effectiveness depends heavily on societal conditions. To maximize the impact of campaigns, it is imperative to customize campaign strategies and target geographic regions and age ranges to align with the societal conditions.

The success of the model's vaccination campaigns hinges on societal characteristics that shape the inherent persuasiveness of pro-vaccination opinions within a society. When initial persuasiveness is high, such as with low perceived risk (Figures 3.3a and 3.3b), natural uptake is already high, limiting the impact of the campaign. When initial persuasiveness is low, such as with high perceived risk (Figure 3.3g), natural uptake is low. Yet, our campaign, which focuses on increasing opinion sharing, is ineffective because the underlying persuasiveness towards vaccination is low. This implies that simple encouragement alone is insufficient if there is low societal trust in vaccination. The campaign is most effective in societies with moderate levels of opinion persuasiveness, such as those with moderate perceived risk and where individuals balance social influence and evidence-based decision-making factors (Figure 3.3e). In this case, stimulating discussions through the campaign significantly boosts vaccination rates by building upon existing persuasive-

ness.

The optimal target demographic for campaigns also depends on the societal characteristics that influence initial level of opinion persuasiveness. To maximize health benefits, campaigns initially prioritize group (Region 1, Age 18-34), as this group has a high potential for opinion changing impact. However, in societies with high opinion persuasiveness (Figures 3.3a and 3.3b), highly populated areas quickly achieve sufficient vaccination coverage due to their strong social interactions. When this occurs, campaign resources shift to rural areas to ensure these regions benefit directly. In scenarios where campaigns are likely to be ineffective due to low persuasiveness (Figure 3.3g), targeting the 35-49 age range is a better strategy. This approach indirectly promotes vaccination among 50 and over age range, who are more persuadable due to their higher mortality risk. This observation is consistent in our one-way sensitivity analysis results.

Limitations exist in our model. Our model relies on several simplifying assumptions regarding how vaccination opinions spread. The model assumes that anti-vaccination individuals only change their opinion through social interactions. In reality, individuals may adjust their views in response to new information or personal experiences. Our model assumes a direct link between vaccination opinions and behavior, in that all pro-vaccination individuals receive a vaccine immediately. In reality, there is a gap between belief and action, due to insufficient vaccine supply or poor access to vaccines. Pro-vaccinators maintain their view during our one-year time horizon. Our grouped compartmental model sacrifices the details of individual heterogeneity. This includes variations in levels of social connections and opinion characteristics that may differ by individuals.

Our vaccination campaign model focuses on increasing the number of pro-vaccination opinion sharers via word-of-mouth effects. We have not tailored messages to the diverse demographics and individual preferences that can impact persuasiveness. However, our findings suggest this is a promising approach, such as emphasizing safety concerns for evidence-based thinkers or community vaccination rates for those who are more socially influenced. We also did not consider direct recruitment incentives as tested by other studies [69, 111]. While we explored communication channels with modifications to the opinion sharing matrix, the campaign primarily relied on word-of-mouth methods. An alternative to this method would be other communication methods, such as broadcast media. This might require adjustments to the force of opinion propagation $\theta_i(x,t)$, where the campaign introduces new influences to stimulate thoughts on vaccination, extending beyond opinion sharing within a social network. To account for this, data on how different demographics interact with these channels will be needed.

The model draws on real-world data on demographics, social networks, and disease and opinion parameters that are calibrated to 2023 COVID-19 epidemic data. Changes in disease characteristics or demographics could change the optimal target groups and campaign approaches. The model is flexible and can accommodate any population grouping (e.g., by political stance, occupation, social interest) provided there is sufficient data. Further, while we use a specific definition of perceived vaccine benefits and risks, alternative definitions could yield different insights, though overall trends may remain similar.

Our two-stage coarse grid search approach effectively identified optimal target groups within the chosen objective function. Our initial comparisons with global optimization approaches suggest consistency in outcomes. However, more sophisticated algorithms could be needed when exploring high-dimensional formulation.

Future research should refine the model in several key areas. First, we should account for individual heterogeneity in social connectedness and decision-making processes through individual-level models. While offering a more detailed model, this approach might require more complex optimization techniques and encounter data-related challenges compared to our network-based compartmental model. Second, opinion dynamics should be improved by exploring models like complex contagion (i.e., requiring multiple exposures) or threshold-based models (i.e., triggering rapid spread at a critical threshold value) [153]. Third, future research should explore details on campaign interventions, including the impact of varied campaign content and channels. Lastly, analyzing the spread of misinformation and fake vaccine effects would provide crucial insights into their role in hesitancy.

In conclusion, our coupled dynamics model highlights the critical need for societal-specific vaccination campaigns. These include understanding perceived vaccine risk and importance of recognizing social influence or evidence-based reasoning. Our findings call for future research delving deeper into individual differences and diverse messaging that can effectively achieve a safer community.

Chapter 4

REPRESENTATIVE CALIBRATION

Model calibration, the process of estimating model parameters with observed data, is a fundamental step across various modeling methodologies. In Chapters 2 and 3, where large-scale agent-based models (ABMs) and compartmental models were employed, considerable effort was dedicated to the calibration procedure. This involved the careful selection of data and the identification of key parameters for calibration. Furthermore, we recognized the importance of determining the optimal number of parameter value sets needed to adequately represent model behavior under real-world uncertainties.

Driven by the challenges of calibration, this chapter addresses research objective (3) proposing a calibration framework that addresses model representativeness under data uncertainties. The two-stage approach focuses on efficiently identifying multiple diverse parameter sets to represent target data. In the first stage, a black-box optimization algorithm generates near-optimal parameter sets. The second stage then clusters these sets, reducing the number of parameter value sets while representing diverse model behaviors. This study has been published in 2023 Winter Simulation Conference [85].

4.1 Introduction

Calibration, or parameter estimation to fit a model to data, is essential for ensuring the validity of a model and model outcomes (e.g., simulation results). Model calibration typically involves four steps: identify the parameters to be calibrated, select target data to compare with model outcomes, determine a goodness-of-fit (GOF) measure between target data and model outcomes, and choose parameter search strategies [134].

Previous research on model calibration considers two main approaches: i) identifying a single optimal parameter set, and ii) determining a large number of feasible parameter sets. Relying solely on a single parameter set may not account for uncertainty in the target data and may limit the range of future model predictions beyond the calibration period. A large number of parameter

sets may be a better representation of uncertainty, but has a high computation cost.

In our previous study [86] (Chapter 2), we developed an agent-based model where we calibrated the model for 152 days and then predicted policy outcomes for 348 days. During the prediction period, we conducted a 4-way sensitivity analysis on policy interventions, resulting in over 300 policy scenarios for each calibration set. Due to the complexity and stochastic nature of the model, each policy run took around 500 CPU minutes on the Hyak supercomputer system at the University of Washington. To mitigate the computational resources needed, we employed a clustering approach, reducing the number of parameter sets to two. This enabled us to obtain valuable policy insights while avoiding excessive use of computer resources.

Inspired by our previous study, we propose a two-stage process that we call representative calibration. This approach aims to identify multiple diverse parameter sets that are both computationally efficient and "good enough" to represent the target data. Our calibration approach involves a two-stage process. In the first stage, we leverage well-known black-box optimization algorithms for parameter search and identify good-enough parameter sets. In the second stage, we apply a clustering approach to the selected parameter sets to obtain representative parameter sets. This two-stage approach balances computation efficiency with sufficient coverage of potential future model outcomes.

In this study, we evaluate the effectiveness of our proposed calibration approach and compare the performance of several parameter search algorithms. In Section 4.2, we review relevant literature on model calibration. Section 4.3 explains the concept of representation calibration, including black-box optimization algorithms and clustering. We describe the numerical comparison plan in Section 4.4.1, and present a disease-opinion compartmental model [84] in Section 4.4.2 to illustrate the two-stage calibration process. Finally, we present the study's findings in Section 4.5 and provide a discussion in Section 4.6.

4.2 Literature Review

In recent years, various calibration techniques have been proposed across different fields such as epidemics, economics, engineering, and neuroscience. As models become more complex, data availability increases, and methodology advances, calibration techniques have become more effective and efficient. Most of these techniques focus on time efficiency and precision of the model to target data, with little attention given to the robustness or representativeness of the calibration.

Latin hypercube sampling has been commonly used in calibration due to its simplicity and coverage of a parameter space [102], and is still in use today ([86, 115]). Black-box optimization methods, such as simulated annealing and genetic algorithms have also been applied to calibration ([34, 40]). Bayesian optimization methods, such as the Sequential Model-based Algorithm Configuration (SMAC) [63] and Optuna [2] have been used for hyperparameter tuning and calibration ([77, 100]). Our study explores these well-known calibration algorithms within the proposed two-stage representative calibration framework.

Several Bayesian calibration methods address parameter uncertainty by approximating a posterior distribution. Kerr et al. [77] used Optuna to calibrate an agent-based epidemiological model. The study derived a posterior distribution of parameters by using more than 15,000 parameter sets. The top ten best-fitting parameter sets were used for scenario analyses. Jalal et al. [66] combined Bayesian calibration with an artificial neural network as a surrogate model. By generating 10,000 parameter sets, they derived a posterior distribution of parameters while accounting for data uncertainty through the use of threshold values. The paper improved accuracy and computation time compared to an importance sampling algorithm.

A Gaussian process metamodel was used in Xie et al. [149] to obtain the posterior distribution of parameters, and quantified a credible interval to account for parameter estimation uncertainty. Unlike our approach which aims to capture uncertainty of the model, the calibration performance was focused on reducing the width of credible intervals and prediction intervals. Nevertheless, it remains unclear whether these approaches represent diverse uncertainty in model outcomes, since the top parameter sets were selected solely based on GOF measure.

A clustering approach was adopted in Krauledat et al. [80] to reduce the calibration process for Brain-Computer Interfaces. The calibration parameters consist of prototypes of a Common Spatial Pattern algorithm used to classify brain states. Although this study uses clustering for calibration, the focus of the study is to reuse previously calibrated clustering from previous data for new data. This is slightly different from our objective in that our main interest lies in efficiently calibrating a model to given data, and then predicting future trajectories, instead of reusing clustered results to new data.

Most of the algorithms focus on time efficiency and precision of the model to target data. The calibrated parameter sets are usually presented as each individual parameter's posterior distribution. Little attention has been given to the information or representativeness of the whole calibration parameter space on the model outcomes.

4.3 Representative Calibration

We propose a two-stage process, where the first stage is to optimize the GOF using an optimization algorithm and the second stage is to apply clustering to the results of the first stage optimization, obtaining representative parameter sets.

In the first stage, we use a black-box optimization algorithm to search for parameters and identify parameter sets that minimize the goodness-of-fit measure between model outcomes and target data. The first stage optimization problem is stated as,

$$\begin{aligned} \min_{x} \quad & \text{GOF}(f(x), y) \\ & l_i \leq x_i \leq u_i \qquad \text{for } i = 1, \ldots, n \end{aligned} \tag{4.1}$$

where the target data y is a vector in m dimensions, $y = [y_1, \ldots y_m]$, and the model outcome, $f(x)$, may also be m-dimensional to correspond to the target data. The calibration parameter vector $x = [x_1, \ldots, x_n]$ denotes a calibration parameter set in n dimensions, and typically has lower and upper limits l_i and u_i, respectively, for $i = 1, \ldots, n$. The goodness-of-fit measure is defined by the user to be appropriate to the model and data (e.g., Mean Square Error, Mean Absolute Error, or Total Sum of Squares).

In the second stage, we apply a clustering technique to the parameter sets obtained from the first-stage to identify representative parameter sets. After identifying the representative parameter sets, the model is run for a period longer than the calibration period to observe the diversity of predicted future model trajectories.

4.3.1 Stage 1: Parameter Search using Black-Box Optimization

This section describes the black-box optimization algorithms used in the first stage of representative calibration. The algorithms include Latin Hypercube Sampling, Sequential Model-based Algorithm Configuration, Optuna, Simulated Annealing, and Genetic Algorithm.

Latin Hypercube Sampling (LHS)

Latin hypercube sampling is a quasi-random sampling method that is often favored in computer experiments because of its simplicity and coverage of the parameter space [102]. LHS divides each parameter into equally probable intervals and samples once from each interval. Such even spacing

of samples reduces sampling variance and can be applied to high-dimensional problems. However, when the sample size is not large enough, LHS may not be as effective as other algorithms at minimizing GOF.

Sequential Model-based Algorithm Configuration (SMAC)

SMAC is a Bayesian optimization framework that is applied in various areas, ranging from hyperparameter tuning in machine learning to global optimization of black-box functions [63]. SMAC creates a random forest surrogate model and updates it as the algorithm proceeds. SMAC combines a local search with random sampling to balance exploration and exploitation [7]. Although SMAC's Bayesian approach may find near-optimal solutions with a small number of model runs, it may take a long time to build the random forest, making SMAC appropriate for computationally expensive models.

Optuna

Optuna is an optimization framework actively used in hyperparameter tuning that uses a dynamic approach to explore the search space [2]. It employs a combination of sampling and pruning algorithms to improve the efficiency of the optimization algorithm. The default sampling method for Optuna uses the Tree-structured Parzen estimator (TPE), which generates two probability density functions for "good" and "bad" subsets [15]. As default, the median pruning technique is used, which terminates a model if its best performance is inferior to the median of all model outcomes [55]. Similar to SMAC, Optuna may require only a small number of model runs, but the additional effort required to build the TPE and median pruning may result in computational overhead.

Simulated Annealing (SA)

Simulated annealing is a metaheuristic global optimization algorithm that randomly samples from a domain and accepts a candidate point based on a "cooling schedule" that gradually decreases over time [103]. This approach helps the algorithm escape local minima and find approximate global optima, which is difficult for other optimization techniques such as gradient descent. However, the performance of simulated annealing depends on the method for generating sequential points and the tuning of the cooling schedule.

Genetic Algorithm (GA)

Genetic algorithms are metaheuristic global optimization algorithms that mimic natural evolution by utilizing the survival of the fittest, selection, and mutation [60]. GA evaluates solutions based on a fitness function, selects the best ones, and generates new populations using genetic operators like crossover and mutation. GA can dynamically change the search process by varying crossover and mutation probabilities, but it can be computationally expensive for complex problems that require large population sizes and high numbers of generations. Results can also be sensitive to the initial population.

4.3.2 Stage 2: Clustering

In the second stage, we reduce the number of parameter sets by clustering a set of good-enough parameter sets that satisfy a GOF threshold. Our method for determining the optimal number of clusters, inspired by the elbow method introduced by [78], involves identifying the value of K at which an additional cluster (i.e., increasing the number of clusters from K to $K+1$) does not lead to a reduction of more than 5% in the total within-cluster sum of squares (WSS).

Next, the Partition Around Medoids (PAM) algorithm [75] is applied to identify K-medoid points as representative parameter sets. Unlike the centroid method that calculates the mean of all points within a cluster, the medoid is the actual point in the cluster that is most centrally located. It is less susceptible to extreme values that may skew the mean in the centroid approach. Additionally, since the medoid is an actual data point, it has a more straightforward interpretation and can be more easily related to the original data. While other clustering methods could be employed, the robustness of PAM to outliers and its focus on representative data points align well with our goal of identifying a representative parameter sets.

4.4 Computational Study

4.4.1 Experimental Setup

Our computational study aims to explore a two-stage representative calibration process with five different first-stage optimization algorithms namely, LHS, SMAC, Optuna, SA and GA. We solve the calibration optimization problem (4.1) with mean absolute error as our goodness-of-fit measure. We run each algorithm for 50,000 function evaluations. For each algorithm, instead of

executing 50,000 function evaluations in one run, we run 10 replications of the algorithm with 5,000 function evaluations in each replication. This scheme, from our observation from numerical experiment, balances exploration with exploitation to some extent. For simulated annealing, our numerical experience is that the solution plateaus after about 5,000 function evaluations. So repeating simulated annealing with 10 different starting points (associated with the initial random seed), we hope to identify more good points. Hence, our experiment runs each algorithm for 50,000 function evaluations total, but is the result of 10 repetitions with 5,000 function evaluations each. After the 50,000 function evaluations, we filter the 50,000 points to obtain those with MAE < 2, which we consider good-enough parameter sets.

In the second stage, we employ the Partition Around Medoids algorithm to identify medoid points that represent clusters of the good-enough parameter sets. We then simulate the model for 100 days (59 days for the calibration period and 41 days beyond the calibration period) with the clustered parameter sets and evaluate the result.

All experiments are performed on a high-performance computing cluster equipped with 10 CPU cores and 20GB of memory, using an Intel Xeon processor.

4.4.2 Model Description

In this section, we define the parameters to be calibrated, the target data, and the goodness-of-fit measure to evaluate the similarity between target data and model outcomes. We illustrate our calibration method using a disease-opinion compartmental model on COVID-19. This example is a previous version of Chapter 3. The calibration approach remains largely similar with the exception that in the newer version, parameters x_7 through x_{13} are fixed rather than calibrated. Moreover, the newer version explicitly sets confidence intervals for each dataset by measuring data trustworthiness from literature. See Appendix B.3 for the updated calibration procedure.

We calibrate thirteen unknown parameters, denoted as $x = (x_1, \ldots, x_{13})$, where each parameter x_i has a lower bound and an upper bound, $l_i \leq x_i \leq u_i$, as presented in Table 4.1. The calibration parameters consist of four disease-related parameters (x_1, x_3, x_4, x_5) and nine opinion-related parameters $(x_2, x_6, \ldots x_{13})$.

Figure 4.1 displays the target data y which consists of eight weekly data points from January to February 2023 and comprises five types of data. We use the notation $y_{k,t}$ to represent the target data, where k refers to the type of health outcome $k \in \{1, \ldots, 5\}$ and t denotes each week during the period $t \in \{1, \ldots, 8\}$. Specifically, $k = 1, 2,$ and 3 correspond to the percentage of the population

Parameter (x_i)	Description	Lower bound (l_i)	Upper bound (u_i)
x_1	Transmission rate of infectious individuals	1.5	2.5
x_2	Average number of opinion sharing contacts per day when sharing opinion	0.1	0.5
x_3	Infection-fatality rate of reference age range (18 to 29 years)	0.0002	0.0003
x_4	Average number of days from recovery to become susceptible	150	250
x_5	Proportion of initially susceptible individuals for all groups	0.4	0.8
x_6	Initial proportion of pro-vaccinators that share vaccination opinion	0.01	0.05
x_7	Emotional judgment importance for vaccination in age range 0-17	0.01	0.99
x_8	Emotional judgment importance for vaccination in age range 18-34	0.01	0.99
x_9	Emotional judgment importance for vaccination in age range 35-49	0.01	0.99
x_{10}	Emotional judgment importance for vaccination in age range 50-64	0.01	0.99
x_{11}	Emotional judgment importance for vaccination in age range 65 and over	0.01	0.99
x_{12}	Sensitivity of rational judgement to vaccination probability	0.1	1.0
x_{13}	Sensitivity of emotional judgement to vaccination probability	0.1	1.0

Table 4.1: 13 calibration parameters (x_i) and their corresponding lower bounds (l_i) and upper bounds (u_i).

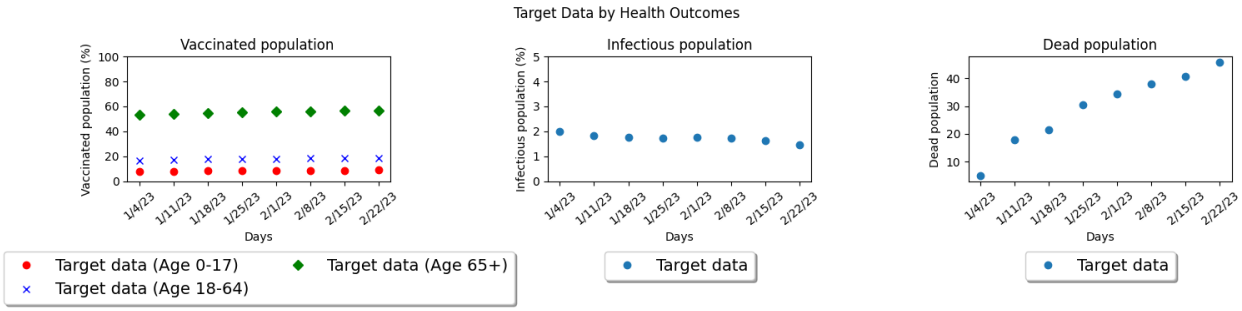


Figure 4.1: Target data consisting of eight weekly data points on vaccination, infectious, and dead population from January to February 2023.

that is vaccinated in age group 0 to 17, 18 to 64, and over 65, respectively; $k = 4$ denotes the percentage of the population that is infectious; and $k = 5$ represents the running cumulative number of deaths. Additionally, we use $f_{k,t}(x)$ to represent the five model health outcomes at the eight weekly dates, that correspond to the calibration parameter set x.

Our GOF measure between target data $y_{k,t}$ and model outcomes $f_{k,t}(x)$ is calculated as the sum of mean absolute errors (MAE), which is then normalized by the average of each target data. We use MAE instead of mean squared error (MSE) because our normalized measure ranges between 0 to 1, so the absolute difference provides an estimate of the absolute error, whereas MSE tends to underestimate larger differences in the 0 to 1 range. The calibration optimization problem is formulated as follows:

$$\begin{aligned} \min_{x} \quad & \text{GOF}(f(x), y) \\ & \text{GOF}(f(x), y) = \sum_{k=1}^{5} \sum_{t=1}^{8} \frac{\left| f_{k,t}(x) - y_{k,t} \right|}{\bar{y}_k} \\ \text{s.t.} \quad & l_i \leq x_i \leq u_i \qquad \text{for } i = 1, \ldots, 13 \end{aligned} \tag{4.2}$$

where $y_{k,t}, f_{k,t}(x) \in \mathbb{R}$ for $k = 1, \ldots, 5$ and $t = 1, \ldots, 8$, and $\bar{y}_k = (1/8) \sum_{t=1}^{8} y_{k,t}$ for $k = 1, \ldots, 5$.

4.5 Results

4.5.1 Parameter Search Results using Black-Box Optimization

This section presents the results from the first stage of representative calibration. As shown in Table 4.2, SMAC, SA, and GA have a higher number of good enough points (MAE < 2) out of 50,000 points (10 runs with 5,000 function evaluations each). This indicates that SMAC, SA, and GA can identify good solutions quickly and more than two-thirds of the solutions found have MAE < 2. The computation time of SMAC and Optuna is about twice that of LHS, SA, and GA, due to the overhead of computing a random forest and TPE, respectively.

Algorithm	LHS	SMAC	OPTUNA	SA	GA
Best MAE	1.613	1.350	1.455	1.346	1.378
Number of good enough points (MAE < 2)	318	37,505	7,331	31,562	47,093
Total Time (seconds)	12,157	29,631	13,479	12,296	26,612

Table 4.2: Performance measures for each of five algorithms including, the best incumbent MAE value of 50,000 points, number of good-enough points (MAE < 2) out of 50,000 points, and total computation time in seconds.

Figure 4.2 shows the incumbent function values averaged over 10 runs of each algorithm, as well as the maximum and minimum value over 10 replications represented by vertical bars at selected numbers of function evaluations. Observe that the result of this plot supports why we did 10 replications of 5,000 function evaluations in each run instead of 50,000 sequential iterations. The plot clearly demonstrates that the algorithms progress in the initial iterations, followed by a slowdown and eventual plateauing, signifying a diminishing return on subsequent iterations. Observe that, in Figure 4.2, SMAC and SA outperform Optuna and GA in terms of minimizing MAE. Table 4.2 also shows that the best incumbent MAE value for SMAC and SA is lower than

the others.

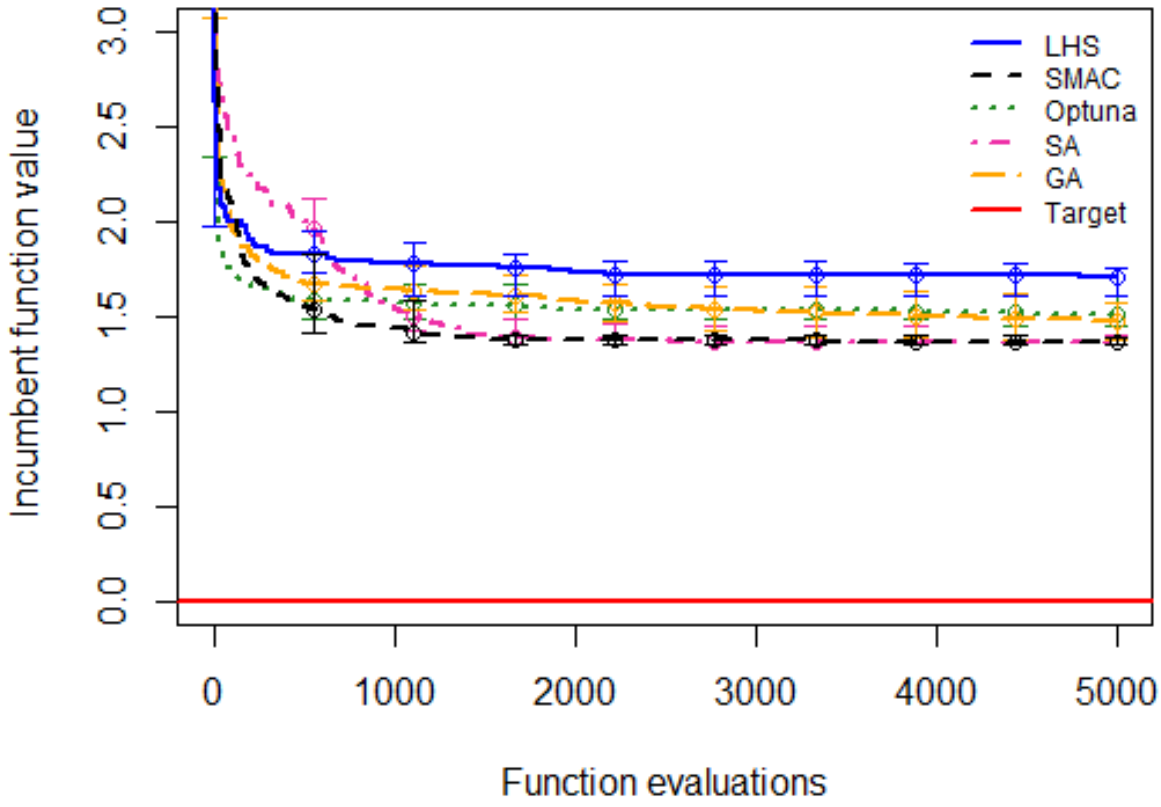
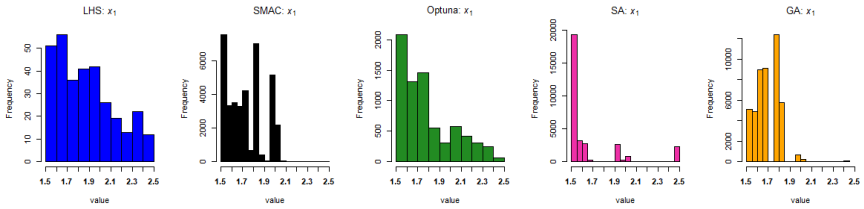


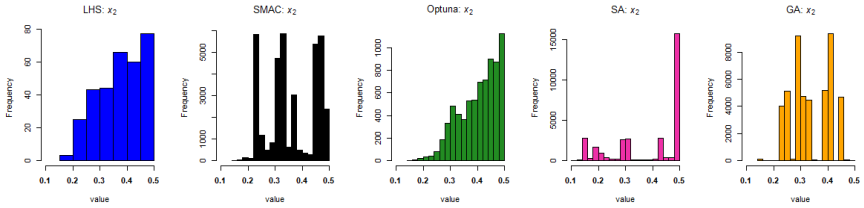
Figure 4.2: Incumbent function value plots for the average, maximum and minimum MAE over 10 runs of the five algorithms (LHS, Optuna, SMAC, SA, and GA) applied to the disease-opinion compartmental model.

Figure 4.3 shows histograms of the two most important parameters, x_1 and x_2, from the filtered solutions (MAE < 2) over 50,000 function evaluations from the five algorithms. The distribution of parameter values within their lower and upper bounds differs by algorithm. LHS provides a spread of possible values over the lower and upper bound range, whereas SA and GA are highly concentrated around the near-optimal values. The parameter values from Optuna are nearly as spread out as LHS, but still concentrate on near-optimal values. The distribution is one aspect of

representative parameter sets.



(a) Histogram of parameter x_1



(b) Histogram of parameter x_2

Figure 4.3: Histograms of parameters x_1 and x_2 illustrate the distribution of good enough solutions (MAE < 2) from all the five algorithms (LHS, Optuna, SMAC, SA, and GA) over all 50,000 function evaluations.

4.5.2 Clustering Results

The Partition Around Medoids algorithm is used to identify K-medoid points as representative parameter sets. For each algorithm, the PAM algorithm is applied to the filtered set of good-enough parameter sets over 50,000 points. We plot the cluster results on the two most important parameters, namely, x_1 and x_2. Figure 4.4 illustrates the clustering of good enough parameter sets from LHS, SMAC, Optuna, SA, and GA. Each dot in the 2-dimension plot represents a good-enough parameter set, and the large circle represents the medoid point of a cluster. Since the number of clusters K is determined by observing changes in the total within-cluster sum of squares, each algorithm has a different number of clusters, as a result, a different number of medoid points, i.e., LHS has 4 clusters, SMAC has 10 clusters, Optuna has 7 clusters, SA has 12 clusters, and GA has 7 clusters, as shown in Figure 4.4. In Figure 4.4, the good-enough points of LHS and Optuna were scattered over both parameter ranges. On the other hand, points from SMAC, SA, and GA focus on a narrower range. Additionally, Optuna was able to provide a higher density of good enough solutions than LHS, with a wide range of representative medoid points. Note that

the clustering results from Figure 4.4 are plotted for two parameters, x_1 and x_2, out of thirteen parameters. The visualization of the nearest medoid in two dimensions does not illustrate the full thirteen dimension space.

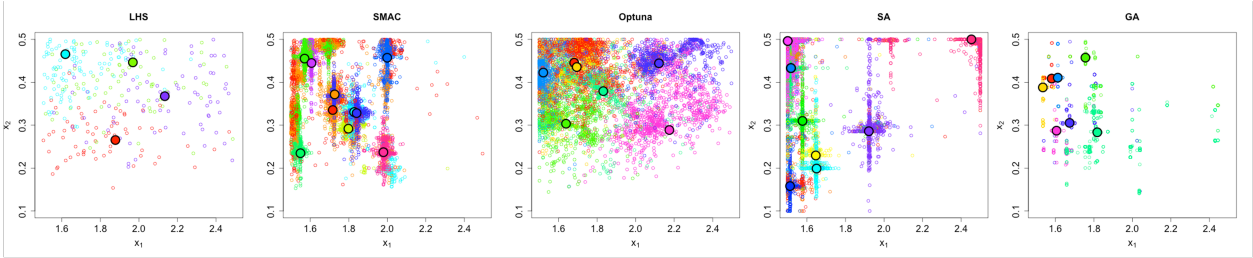
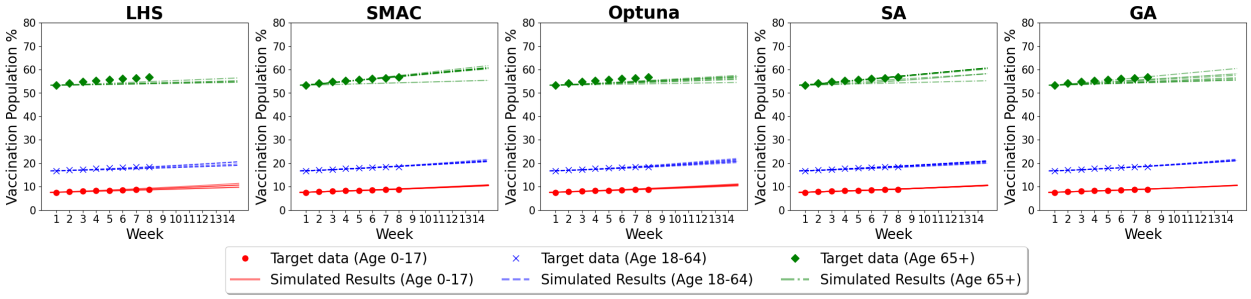


Figure 4.4: Cluster results on parameters x_1 and x_2 out of 13 parameters, with medoid points for each algorithm.

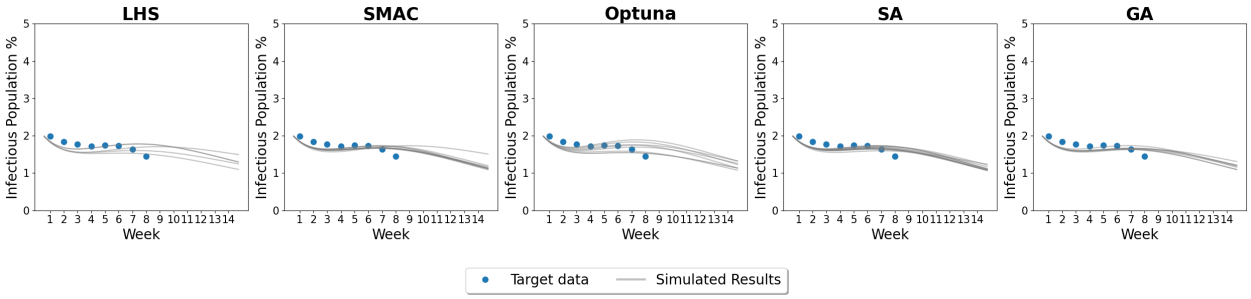
4.5.3 Model Trajectories for 100 Days

Finally, for each black-box optimization algorithm, we run the disease-opinion compartmental model, using the K-medoid parameter sets for 100 days (59 days used for calibration and 41 days for prediction). We then evaluate the diversity of model trajectories by plotting health outcomes for each representative parameter set. Figure 5 presents the five health outcomes from the compartmental model: the vaccinated population by age group (Figure 4.5a), infectious population (Figure 4.5b), and dead population (Figure 4.5c). In each plot, a single line represents the model trajectory using a medoid point for the parameter set. As such, the number of lines corresponds to the number of clusters from each algorithm. Therefore, LHS, SMAC, Optuna, SA, and GA have 4, 10, 7, 12, and 7 lines, respectively.

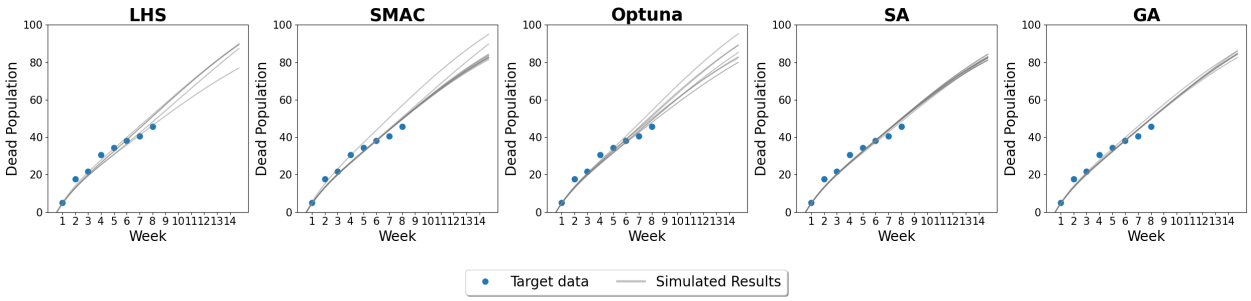
LHS (with 4 medoid points) and Optuna (with 7 medoid points) show diverse results in infectious population and dead population from the span of the trajectories. The trajectories for the vaccinated population in the 65+ age group match the target data and are more diverse with SMAC, SA, and GA than for Optuna and LHS. This is helpful in representing the uncertainty of the target data and impact on health outcomes. On the other hand, SMAC, SA, and GA, with multiple concentrated medoid points (10, 12, and 7, respectively) tend to show trajectories that lie close to each other.



(a) Vaccinated population



(b) Infectious population



(c) Dead population

Figure 4.5: Health outcomes from the disease-opinion compartmental model, (a) vaccinated population, (b) infectious population and (c) dead population, using the medoid parameter sets for all five algorithms, LHS, Optuna, SMAC, SA, and GA respectively (from left to right).

4.6 Discussion

In this paper, we evaluate a representative calibration approach that identifies multiple diverse parameter sets in order to represent the uncertainty of predicted model outcomes while being computationally efficient. The approach involves a two-stage process, (1) apply black-box optimization algorithms to search for good parameter sets, then (2) apply a clustering approach to obtain repre-

sentative parameter sets.

We observe that in the first stage, most algorithms find good-enough solutions that have relatively low MAE. Considering the best MAE value discovered, SA achieves the lowest MAE, followed by SMAC, GA, Optuna, and LHS, respectively. SMAC converges to a good value faster in the early iterations, as the algorithm uses a random forest surrogate model and a Bayesian approach combining a local search and random sampling. This allows SMAC to find optimal solutions with comparatively few function evaluations. However, SMAC requires the most computation time as shown in Table 4.2.

In the cluster analysis from stage 2, each black-box optimization algorithm provides slightly different cluster formations. The LHS algorithm, which divides each parameter into equally probable intervals, shows more variety in the solutions. Hence, the solutions from LHS are scattered over the parameter range and are not concentrated on near-optimal areas. On the other hand, SMAC and SA, which perform the best in terms of minimizing MAE, provide multiple clusters of solutions, however, the location of each cluster is in a narrower range of solutions. Optuna also allows more exploration, as evidenced by the histogram and cluster results.

While SMAC and SA achieved the smallest GOF relatively quickly, we do not necessarily consider the algorithms as the best calibration method. Our rationale is that the target data has uncertainty so achieving the sole minimization of GOF may be less crucial for robust modeling. Instead, we aim to identify a broad range of good-enough parameter sets, as this is more important in enhancing model robustness. Thus, we find the results from LHS and Optuna more aligned with our goal, as they offer a wider variety of parameter sets and future trajectories.

We attribute the differences in algorithmic outcomes to the strategies employed for exploration and exploitation. SMAC and SA excel at exploiting and finding optimal solutions, whereas LHS and Optuna excel at exploring and uncovering a wide range of good-enough solutions. We recommend that modelers explore more regions once good-enough solutions are found, which can be achieved through strategies such as having enough different starting points instead of a single long-run, or adjusting algorithmic parameter settings to explore other regions once a threshold is met.

Another consideration for calibration is the choice of GOF measures. When the target data consists of multiple types, such as our model's target data type (i.e, vaccination rates across different age groups, infectious and dead population), the GOF measure may be aggregated to a single value. In our study, we assigned equal weights to each target data type. As shown in Figure 4.5a,

the calibration trajectories exhibit a closer match with the red and blue lines, while relatively less alignment with the green line. Hence, we recommend modelers to carefully select appropriate GOF aggregate metrics based on their specific context and priorities. While we aggregated the GOF scores and selected a threshold value (i.e., MAE < 2), when individual target data have different implications, importance, or reliability, it may be advisable to set different threshold values for each target data type. We will pursue this avenue in the future.

One reason for clustering is to reduce the computation time when performing model analysis. This is especially beneficial in cases where the computation time of the model after the calibration period is longer than the computation time during the calibration period. As an example, in [86], the model runtime for the calibration period (152 days) was approximately 500 minutes, whereas the runtime for the prediction period (348 days) took over twice that. Additionally, the model was used to analyze 300 policy scenarios over the prediction period. In this type of situation, it is desirable that a few selected parameter sets are representative to reflect uncertainty in the target data and corresponding model outcomes.

Limitations exist in our model. While we constrained each calibration parameter with box constraints (l_i and u_i), this may limit the parameter space. As the choice of these bounds may impact the model performance, careful consideration is needed in choosing the bounds. While we chose the PAM algorithm to identify K-medoid point, other clustering algorithms should be explored.

In summary, we propose that modelers consider parameter calibration from various perspectives. Instead of solely aiming to minimize the goodness-of-fit quickly, they should 1) recognize that data inaccuracies may exist and allow a certain level of error between the data and model outcomes, 2) formulate the calibration problem with GOF and threshold to identify a range of good-enough parameter sets, 3) explore a broad range of parameter sets so that the model captures the variability of possible future trajectories, and 4) find a sweet spot between exploration and exploitation, wherein exploration promotes broad range of parameter sets while exploitation enhances accuracy in aligning with the target data.

Chapter 5

CONCLUSION AND FUTURE WORK

This dissertation research addresses critical public health problems by applying and advancing healthcare policy modeling approaches. It develops methods that incorporated complex individual behaviors and addresses uncertainties inherent in data and simulations. The dissertation focuses on challenges such as identifying optimal NPIs and vaccination policies during COVID-19 pandemic and tackling vaccine hesitancy through targeted vaccination promotion campaigns. To enhance model generalizability and account for data uncertainty, a calibration framework is developed to identify minimal sets of parameter values that capture a diverse range of model behaviors.

During the COVID-19 pandemic, policymakers were in rapid need to identify effective public health interventions that balance safety with minimal social disruption. In Chapter 2, agent-based simulation models are developed to identify optimal NPI and vaccination policies. This model's ability to directly model public health policies such as face mask use and social distancing enable explicit interpretation of their impact on disease outcomes. In collaboration with local public health experts, this was especially valuable during the early pandemic when data was limited. This research investigates policy impacts under parameter uncertainties by employing either representative calibration or simulation of a wide range of viral mutations. By analyzing the results, the studies provide insights that can guide policy decisions across diverse epidemic scenarios, demonstrating the adaptability and practical value of modeling for pandemic response.

While this research offers valuable contributions, there is potential to advance the agent-based modeling in public health. Instead of assuming individuals' steady compliance to public health policies, future studies might benefit from considering dynamic human behavior, where individuals' responses to policies may evolve over time. Future research could also apply and test the model in diverse regional settings, exploring how variations in the demographics and social networks might influence policy impacts. In order to provide robust policy analysis, methods for addressing the inherent stochasticity of agent-based models, particularly in the early phases of an outbreak, should be considered.

Although vaccines have been proven effective in preventing infections or reducing disease

severity, vaccine hesitancy remains a persistent public health challenge. Chapter 3 tackles this challenge by modeling the coupled dynamics of disease transmission and vaccination opinion propagation, and the impact of word-of-mouth vaccination campaigns. This research employs an evolutionary game-theoretic framework to model dynamic human behavior, overcoming limitations in my earlier work of static behavior. Leveraging real-world data on population demographics and social connections, our findings offer valuable insights for policymakers at the population level. A key finding is that optimal targeting and messaging strategies must be tailored to specific demographic groups within a community. This approach acknowledges that individuals' vaccination behaviors are shaped by both evidence-based factors (like risk-benefit assessment) and social influence factors (such as the vaccination decisions of their social networks). By developing a flexible model that accommodates various population groupings (e.g., by political stance, occupation, social interest), this research provides actionable guidance applicable to diverse real-world scenarios. In the face of growing vaccine hesitancy, this research is particularly valuable.

To enhance the model's practical applications in real-world policy decisions, future research could refine several parts. This includes modeling individual-level heterogeneity in social connections and decision-making, exploring alternative opinion propagation models (e.g., complex contagion, threshold-based models), and analyzing the impact of misinformation. Research could also analyze the effectiveness of diverse messaging content and communication channels for targeted vaccination campaigns. Refining the policy identification approach by developing and integrating constrained nonlinear optimization algorithms will be crucial as a coarse-grid approach may be insufficient to handle the added complexities within the model.

In complex simulation models, achieving reliable outcomes while maintaining computational efficiency is an ideal goal. Well-designed calibration practices are essential to ensure this. In Chapter 4, I propose a two-stage approach where black-box optimization algorithms and clustering are applied to efficiently select representative parameter subsets. Unlike methods focused solely on speed or precision, this approach prioritizes representativeness, capturing the model's full range of potential outcomes and enhancing its generalizability.

The research goal is to provide a multifaceted calibration framework that balances speed, accuracy, and representativeness. This work offers an initial step, but identify two key areas for future research. First, a more concrete method is needed to define "representativeness" of selected parameter sets. Borrowing diversity and representativeness metrics from Emmerich et al. [47] is a good starting point. Computing the hyperplane of model outcomes from the selected parameter

sets is another potential approach. While their metrics offer value, they do not account for individual parameter influence on model outcomes. Future work could incorporate parameter distances weighted by their importance to model behavior, which might be determined through statistical modeling. Principal component analysis could be applied to evaluate the parameter importance. Additionally, efficient algorithms for selecting representative parameter subsets are necessary. Possible strategies include greedy approaches (starting with the best-fitting set and adding points for representativeness), clustering (using elbow methods to determine optimal clusters and selecting representative points), and regression trees to classify parameter sets.

This dissertation advances healthcare policy modeling by addressing heterogeneous individual behaviors and uncertainties arising from data and simulation complexities. Ultimately, this research provides healthcare policymakers with the tools to make informed, data-driven decisions, enhancing their capacity for effective policy making.

BIBLIOGRAPHY

1. A FAIR Health White Paper MM. Risk Factors for COVID-19 Mortality among Privately Insured Patients: A Claims Data Analysis. Generic. [Online; accessed 10-August-2023]. 2020. Available from: https://s3.amazonaws.com/media2.fairhealth.org/whitepaper/asset/Risk%20Factors%20for%20COVID-19%20Mortality%20among%20Privately%20Insured%20Patients%20-%20A%20Claims%20Data%20Analysis%20-%20A%20FAIR%20Health%20White%20Paper.pdf

2. Akiba T, Sano S, Yanase T, Ohta T, and Koyama M. Optuna: A Next-Generation Hyperparameter Optimization Framework. *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. KDD '19. Anchorage, AK, USA: Association for Computing Machinery, 2019 :2623–31. DOI: 10.1145/3292500.3330701

3. Aleta A, Martín-Corral D, Pastore y Piontti A, Ajelli M, Litvinova M, Chinazzi M, et al. Modelling the impact of testing, contact tracing and household quarantine on second waves of COVID-19. Nature Human Behaviour 2020; 4:964–71. DOI: 10.1038/s41562-020-0931-9

4. Altarawneh HN, Chemaitelly H, Ayoub HH, Tang P, Hasan MR, Yassine HM, et al. Effects of Previous Infection and Vaccination on Symptomatic Omicron Infections. N Engl J Med 2022; 387:21–34. DOI: 10.1056/NEJMoa2203965

5. Altarawneh HN, Chemaitelly H, Hasan MR, Ayoub HH, Qassim S, AlMukdad S, et al. Protection against the Omicron Variant from Previous SARS-CoV-2 Infection. N Engl J Med 2022; 386:1288–90. DOI: 10.1056/NEJMc2200133

6. Altshuler Y, Shmueli E, Zyskind G, Lederman O, Oliver N, and Pentland A. Campaign Optimization Through Behavioral Modeling and Mobile Network Analysis. IEEE Transactions on Computational Social Systems 2014; 1:121–34. DOI: 10.1109/TCSS.2014.2377831

7. Anastacio M and Hoos H. Model-Based Algorithm Configuration with Default-Guided Probabilistic Sampling. *Parallel Problem Solving from Nature – PPSN XVI*. Ed. by Bäck

T, Preuss M, Deutz A, Wang H, Doerr C, Emmerich M, et al. Cham: Springer International Publishing, 2020 :95–110

8. Anderson RM, Heesterbeek H, Klinkenberg D, and Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The lancet 2020; 395:931–4

9. Antia R and Halloran ME. Transition to endemicity: Understanding COVID-19. Immunity 2021; 54:2172–6. DOI: 10.1016/j.immuni.2021.09.019

10. Bailey M, Cao R, Kuchler T, Stroebel J, and Wong A. Social Connectedness: Measurement, Determinants, and Effects. Journal of Economic Perspectives 2018 Aug; 32:259–80. DOI: 10.1257/jep.32.3.259

11. Barber RM, Sorensen RJD, Pigott DM, Bisignano C, Carter A, Amlag JO, et al. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. The Lancet 2022; 399:2351–80. DOI: 10.1016/S0140-6736(22)00484-6

12. Basu A. Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States: Study estimates the COVID-19 infection fatality rate at the US county level. Health Affairs 2020 :10–1377

13. Bauch CT. Imitation dynamics predict vaccinating behaviour. Proceedings of the Royal Society B: Biological Sciences 2005; 272:1669–75. DOI: 10.1098/rspb.2005.3153

14. Bellon T, Baertlein L, and Brown NP. Insight: Personal touch, word of mouth: How U.S. rural communities succeed getting COVID-19 shots into arms. [Online; accessed 17-January-2024]. 2021. Available from: https://www.reuters.com/world/us/personal-touch-word-mouth-how-us-rural-communities-succeed-getting-covid-19-2021-01-29/

15. Bergstra J, Bardenet R, Bengio Y, and Kégl B. Algorithms for Hyper-Parameter Optimization. *Advances in Neural Information Processing Systems*. Ed. by Shawe-Taylor J, Zemel R, Bartlett P, Pereira F, and Weinberger K. Vol. 24. Neural Information Processing Systems. Curran Associates, Inc., 2011

16. Bertsimas D, Boussioux L, Cory-Wright R, Delarue A, Digalakis V, Jacquillat A, et al. From predictions to prescriptions: A data-driven response to COVID-19. Health Care Management Science 2021; 24:253–72. DOI: 10.1007/s10729-020-09542-0

17. Bobrovitz N, Ware H, Ma X, Li Z, Hosseini R, Cao C, et al. Protective effectiveness of previous SARS-CoV-2 infection and hybrid immunity against the omicron variant and severe disease: a systematic review and meta-regression. Lancet Infect Dis 2023; 23:556–67. DOI: 10.1016/S1473-3099(22)00801-5

18. Borges ME, Ferreira LS, Poloni S, Bagattini AM, Franco C, Rosa MQM da, et al. Modelling the impact of school reopening and contact tracing strategies on Covid-19 dynamics in different epidemiologic settings in Brazil. Glob Epidemiol 2022; 4:100094. DOI: 10.1016/j.gloepi.2022.100094

19. Buckley J. Unlocking the World: EU Digital Covid Certificate: Everything you need to know. Generic. [Online; accessed 15-October-2021]. 2021. Available from: https://www.cnn.com/travel/article/eu-covid-certificate-travel-explainer/index.html

20. Canales K. 40% of kids under 13 already use Instagram and some are experiencing abuse and sexual solicitation, a report finds, as the tech giant considers building an Instagram app for kids. [Online; accessed 6-March-2024]. 2021. Available from: https://www.businessinsider.com/kids-under-13-use-facebook-instagram-2021-5

21. Caro JJ, Briggs AH, Siebert U, and Kuntz KM. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Medical Decision Making 2012; 32:667–77

22. Centers for Disease Control and Prevention. COVID-19 Mathematical Modeling. [Online; accessed 22-October-2020]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/covid-data/mathematical-modeling.html

23. Centers for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. [Online; accessed 15-April-2023]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html

24. Centers for Disease Control and Prevention. COVID-19 vaccinations in the United States, County. [Online; accessed 24-March-2023]. 2023. Available from: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh

25. Centers for Disease Control and Prevention. COVID-19: When to Quarantine. [Online; accessed 10-August-2020]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html

26. Centers for Disease Control and Prevention. SARS-CoV-2 Variant Classifications and Definitions. Generic. [Online; accessed 19-March-2021]. 2021. Available from: https://www.cdc.gov/coronavirus/2019-ncov/variants/variant-info.html#Concern

27. Centers for Disease Control and Prevention. Science Brief: SARS-CoV-2 Infection-induced and Vaccine-induced Immunity. Web Page. [Online; accessed 19-October-2021]. 2021. Available from: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html

28. Centers for Disease Control and Prevention. Stats of the State of Washington. Generic. [Online; accessed 19-March-2021]. 2018. Available from: https://www.cdc.gov/nchs/pressroom/states/washington/washington.htm

29. Centers for Disease Control and Prevention. When You've Been Fully Vaccinated: How to Protect Yourself and Others. Generic. [Online; accessed 19-March-2021]. 2021. Available from: https://www.cdc.gov/coronavirus/2019-ncov/vaccines/fully-vaccinated.html

30. Cerqueira-Silva T, Araujo Oliveira V de, Paixao ES, Florentino PTV, Penna GO, Pearce N, et al. Vaccination plus previous infection: protection during the omicron wave in Brazil. Lancet Infect Dis 2022; 22:945–6. DOI: 10.1016/S1473-3099(22)00288-2

31. Chalkias S, Harper C, Vrbicky K, Walsh SR, Essink B, Brosz A, et al. A Bivalent Omicron-Containing Booster Vaccine against Covid-19. N Engl J Med 2022; 387:1279–91. DOI: 10.1056/NEJMoa2208343

32. Champion VL, Skinner CS, et al. The health belief model. Health behavior and health education: Theory, research, and practice 2008; 4:45–65

33. Chemaitelly H, Ayoub HH, Tang P, Coyle P, Yassine HM, Al Thani AA, et al. Long-term COVID-19 booster effectiveness by infection history and clinical vulnerability and immune imprinting. MedRxiv 2022 :2022.11.14.22282103. DOI: 10.1101/2022.11.14.22282103

34. Cheng CT, Zhao MY, Chau K, and Wu XY. Using Genetic Algorithm and Topsis for Xinanjiang Model Calibration With a Single Procedure. Journal of Hydrology 2006; 316:129–40. DOI: https://doi.org/10.1016/j.jhydrol.2005.04.022

35. Choi B, Choudhary MC, Regan J, Sparks JA, Padera RF, Qiu X, et al. Persistence and Evolution of SARS-CoV-2 in an Immunocompromised Host. N Engl J Med 2020; 383:2291–3. DOI: 10.1056/NEJMc2031364

36. Chu DK, Akl EA, Duda S, Solo K, Yaacoub S, Schünemann HJ, et al. Physical distancing, face masks, and eye protection to prevent person-to-person transmission of SARS-CoV-2 and COVID-19: a systematic review and meta-analysis. The lancet 2020; 395:1973–87

37. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HH, Mercer SW, et al. How many are at increased risk of severe COVID-19 disease? Rapid global, regional and national estimates for 2020. MedRxiv 2020. DOI: 10.1101/2020.04.18.20064774

38. Collier N and North M. Parallel agent-based simulation with Repast for High Performance Computing. SIMULATION 2013; 89:1215–35. DOI: 10.1177/0037549712462620

39. Covid-19 Simulator Consortium. Covid19sim - Outbreak Tool Methodology. [Online; accessed 3-October-2020]. 2020. Available from: https://covid19sim.org/

40. Dahabreh IJ, Chan JA, Earley A, Moorthy D, Avendano EE, Trikalinos TA, et al. Modeling and Simulation in the Context of Health Technology Assessment: Review of Existing Guidance, Future Research Needs, and Validity Assessment. Tech. rep. 16(17)-EHC020-EF. Rockville, MD: Tufts Evidence-based Practice Center, 2017. Available from: http://europepmc.org/books/NBK424024

41. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Pearson CA, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine 2020; 26:1205–11. DOI: 10.1038/s41591-020-0962-9

42. Durham DP and Casman EA. Incorporating individual health-protective decisions into disease transmission models: a mathematical framework. Journal of The Royal Society Interface 2012; 9:562–70. DOI: 10.1098/rsif.2011.0325

43. Dyson L, Hill EM, Moore S, Curran-Sebastian J, Tildesley MJ, Lythgoe KA, et al. Possible future waves of SARS-CoV-2 infection generated by variants of concern with a range of characteristics. Nat Commun 2021; 12:5730. DOI: 10.1038/s41467-021-25915-7

44. Earnest R, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, et al. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA. Cell Reports Medicine 2022; 3:100583. DOI: https://doi.org/10.1016/j.xcrm.2022.100583

45. Ehsanfar A and Mansouri M. Incentivizing the dissemination of truth versus fake news in social networks. *2017 12th System of Systems Engineering Conference (SoSE)*. 2017 :1–6. DOI: 10.1109/SYSOSE.2017.7994981

46. Eikenberry SE, Mancuso M, Iboi E, Phan T, Eikenberry K, Kuang Y, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infectious Disease Modelling 2020; 5:293–308. DOI: https://doi.org/10.1016/j.idm.2020.04.001

47. Emmerich MTM, Deutz AH, and Kruisselbrink JW. On Quality Indicators for Black-Box Level Set Approximation. *EVOLVE- A Bridge between Probability, Set Oriented Numerics and Evolutionary Computation*. Ed. by Tantar E, Tantar AA, Bouvry P, Del Moral P, Legrand P, Coello Coello CA, et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013 :157–85. DOI: 10.1007/978-3-642-32726-1_4

48. Ferdinands JM, Rao S, Dixon BE, Mitchell PK, DeSilva MB, Irving SA, et al. Waning 2-dose and 3-dose effectiveness of mRNA vaccines against COVID-19–associated emergency department and urgent care encounters and hospitalizations among adults during periods of Delta and Omicron variant predominance—VISION Network, 10 states, August 2021–January 2022. Morbidity and Mortality Weekly Report 2022; 71:255. Available from: https://www.cdc.gov/mmwr/volumes/71/wr/pdfs/mm7107e2-H.pdf

49. Fisman DN and Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada. CMAJ 2021; 193:E1619–E1625. DOI: 10.1503/cmaj.211248

50. Flagg LA and Anderson RN. National Vital Statistics Reports. Generic. [Online; accessed 19-March-2021]. 2021. Available from: https://www.cdc.gov/nchs/data/nvsr/nvsr69/nvsr69-14-508.pdf

51. Fort H. A very simple model to account for the rapid rise of the alpha variant of SARS-CoV-2 in several countries and the world. Virus Research 2021; 304:198531

52. Fowlkes AL, Yoon SK, Lutrick K, Gwynn L, Burns J, Grant L, et al. Effectiveness of 2-dose BNT162b2 (Pfizer BioNTech) mRNA vaccine in preventing SARS-CoV-2 Infection among children aged 5–11 years and adolescents aged 12–15 years—PROTECT Cohort, July 2021–February 2022. Morbidity and Mortality Weekly Report 2022; 71:422. Available from: http://dx.doi.org/10.15585/mmwr.mm7111e1external

53. Gatto M, Bertuzzo E, Mari L, Miccoli S, Carraro L, Casagrandi R, et al. Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. Proceedings of the National Academy of Sciences 2020; 117:10484–91. DOI: 10.1073/pnas.2004978117

54. Goldberg Y, Mandel M, Bar-On YM, Bodenheimer O, Freedman LS, Ash N, et al. Protection and Waning of Natural and Hybrid Immunity to SARS-CoV-2. N Engl J Med 2022; 386:2201–12. DOI: 10.1056/NEJMoa2118946

55. Golovin D, Solnik B, Moitra S, Kochanski G, Karro J, and Sculley D. Google Vizier: A Service for Black-Box Optimization. *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*. KDD '17. Halifax, NS, Canada: Association for Computing Machinery, 2017 :1487–95. DOI: 10.1145/3097983.3098043

56. Grefenstette JJ, Brown ST, Rosenfeld R, DePasse J, Stone NT, Cooley PC, et al. FRED (A Framework for Reconstructing Epidemic Dynamics): an open-source software system for modeling infectious diseases and control strategies using census-based populations. BMC public health 2013; 13:940. DOI: 10.1186/1471-2458-13-940

57. Gruell H, Vanshylla K, Weber T, Barnes CO, Kreer C, and Klein F. Antibody-mediated neutralization of SARS-CoV-2. Immunity 2022; 55:925–44. DOI: 10.1016/j.immuni.2022.05.005

58. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature medicine 2020; 26:672–5

59. Ho TY, Zabinsky ZB, Fishman PA, and Liu S. Prevention of seasonal influenza outbreak via healthcare insurance. IISE Transactions on Healthcare Systems Engineering 2023; 13:261–78. DOI: 10.1080/24725579.2022.2145393

60. Holland JH. Genetic Algorithms. Scientific American 1992; 267:66–73. Available from: http://www.jstor.org/stable/24939139 [Accessed on: 2024 Apr 18]

61. Holmdahl I and Buckee C. Wrong but Useful — What Covid-19 Epidemiologic Models Can and Cannot Tell Us. The New England Journal of Medicine 2020; 383:303–5

62. Holmes K, Bertozzi S, Bloom B, and Jha P. Disease Control Priorities, Third Edition (Volume 6): Major Infectious Diseases. Disease Control Priorities. World Bank Publications, 2017

63. Hutter F, Hoos HH, and Leyton-Brown K. Sequential Model-Based Optimization for General Algorithm Configuration. *Learning and Intelligent Optimization*. Ed. by Coello CAC. Learning and Intelliigent Optimization (LION). Berlin, Heidelberg: Springer Berlin Heidelberg, 2011 :507–23

64. Institute for Health Metrics and Evaluation. COVID-19 model update: Omicron and waning immunity. Generic. [Online; accessed 19-October-2021]. 2021. Available from: https://www.healthdata.org/special-analysis/omicron-and-waning-immunity

65. Institute for Health Metrics and Evaluation. COVID-19 Projections. [Online; accessed 10-January-2023]. 2020. Available from: https://covid19.healthdata.org/projections

66. Jalal H, Trikalinos TA, and Alarid-Escudero F. BayCANN: Streamlining Bayesian Calibration With Artificial Neural Network Metamodeling. Frontiers in Physiology 2021; 12. DOI: 10.3389/fphys.2021.662314

67. Jalali N, Brustad HK, Frigessi A, MacDonald EA, Meijerink H, Feruglio SL, et al. Increased household transmission and immune escape of the SARS-CoV-2 Omicron variant compared to the Delta variant: evidence from Norwegian contact tracing and vaccination data. MedRxiv 2022

68. Kabir KA and Tanimoto J. Dynamical behaviors for vaccination can suppress infectious disease – A game theoretical approach. Chaos, Solitons & Fractals 2019; 123:229–39. DOI: https://doi.org/10.1016/j.chaos.2019.04.010

69. Kandhway K and Kuri J. Campaigning in Heterogeneous Social Networks: Optimal Control of SI Information Epidemics. IEEE/ACM Transactions on Networking 2016; 24:383–96. DOI: 10.1109/TNET.2014.2361801

70. Kang Y, Gao S, Liang Y, Li M, Rao J, and Kruse J. Multiscale dynamic human mobility flow dataset in the US during the COVID-19 epidemic. Scientific data 2020; 7:390

71. Kaplan RM and Milstein A. Influence of a COVID-19 vaccine's effectiveness and safety profile on vaccination acceptance. Proceedings of the National Academy of Sciences 2021; 118

72. Karimi E, Schmitt K, and Akgunduz A. Effect of individual protective behaviors on influenza transmission: an agent-based model. Health Care Management Science 2015; 18:318–33. DOI: 10.1007/s10729-014-9310-2

73. Kasaie P and Kelton WD. Simulation optimization for allocation of epidemic-control resources. IIE Transactions on Healthcare Systems Engineering 2013; 3:78–93. DOI: 10.1080/19488300.2013.788102

74. Katul GG, Mrad A, Bonetti S, Manoli G, and Parolari AJ. Global convergence of COVID-19 basic reproduction number and estimation from early-time SIR dynamics. PLOS ONE 2020 Sep; 15:1–22. DOI: 10.1371/journal.pone.0239800

75. Kaufman L. Partitioning around medoids (program pam). Finding groups in data 1990; 344:68–125. DOI: https://doi.org/10.1002/9780470316801.ch2

76. Kerr CC, Stuart RM, Mistry D, Abeysuriya RG, Rosenfeld K, Hart GR, et al. Covasim: an agent-based model of COVID-19 dynamics and interventions. PLOS Computational Biology 2021; 17:e1009149

77. Kerr CC, Mistry D, Stuart RM, Rosenfeld K, Hart GR, Núñez RC, et al. Controlling COVID-19 via test-trace-quarantine. Nature Communications 2021; 12:2993. DOI: 10.1038/s41467-021-23276-9

78. Ketchen DJ and Shook CL. The Application of Cluster Analysis in Strategic Management Research: An Analysis and Critique. Strategic Management Journal 1996; 17:441–58. DOI: https://doi.org/10.1002/(SICI)1097-0266(199606)17:6<441::AID-SMJ819>3.0.CO;2-G

79. King County - Department of Health. Summary of COVID vaccination among King County residents. Web Page. [Online; accessed 1-June-2021]. 2021. Available from: https://kingcounty.gov/depts/health/covid-19/data/vaccination.aspx

80. Krauledat M, Schröder M, Blankertz B, and Müller KR. Reducing Calibration Time For Brain-Computer Interfaces: A Clustering Approach. *Advances in Neural Information Processing Systems*. Ed. by Schölkopf B, Platt J, and Hoffman T. Vol. 19. Neural Information Processing Systems. MIT Press, 2006

81. Kucharski A, Klepac P, Conlan A, Kissler S, Tang M, Fry H, et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. The Lancet Infectious Diseases 2020 Jun; 20. DOI: 10.1016/S1473-3099(20)30457-6

82. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine 2020; 172:577–82

83. Le Rutte EA, Shattock AJ, Chitnis N, Kelly SL, and Penny MA. Modelling the impact of Omicron and emerging variants on SARS-CoV-2 transmission and public health burden. Commun Med (Lond) 2022; 2:93. DOI: 10.1038/s43856-022-00154-z

84. Lee S, Liu S, and Zabinsky ZB. Optimizing Vaccination Campaign Strategies Considering Societal Characteristics. Under review at Health Care Management Science. 2024

85. Lee S, Maneekul P, and Zabinsky ZB. Representative Calibration Using Black-Box Optimization and Clustering. *2023 Winter Simulation Conference (WSC)*. 2023 :3669–80. DOI: 10.1109/WSC60868.2023.10408638

86. Lee S, Zabinsky ZB, Wasserheit JN, Kofsky SM, and Liu S. COVID-19 Pandemic Response Simulation in a Large City: Impact of Nonpharmaceutical Interventions on Reopening Society. Med Decis Making 2021; 41:419–29. DOI: 10.1177/0272989X211003081

87. Lee S, Zabinsky ZB, Wasserheit JN, Ross JM, Chen S, and Liu S. Impact of Vaccination and Nonpharmaceutical Interventions With Possible COVID-19 Viral Evolutions Using an Agent-Based Simulation. AJPM Focus 2024; 3:100155. DOI: https://doi.org/10.1016/j.focus.2023.100155

88. Leigh JP, Moss SJ, White TM, Picchio CA, Rabin KH, Ratzan SC, et al. Factors affecting COVID-19 vaccine hesitancy among healthcare providers in 23 countries. Vaccine 2022; 40:4081–9. DOI: https://doi.org/10.1016/j.vaccine.2022.04.097

89. Lemaitre JC, Pasetto D, Zanon M, Bertuzzo E, Mari L, Miccoli S, et al. Optimal control of the spatial allocation of COVID-19 vaccines: Italy as a case study. PLOS Computational Biology 2022 Jul; 18:1–20. DOI: 10.1371/journal.pcbi.1010237

90. Li J, Giabbanelli P, et al. Returning to a normal life via COVID-19 vaccines in the United States: a large-scale Agent-Based simulation study. JMIR medical informatics 2021; 9:e27419. DOI: 10.2196/27419

91. Lin DY, Xu Y, Gu Y, Zeng D, Wheeler B, Young H, et al. Effectiveness of Bivalent Boosters against Severe Omicron Infection. N Engl J Med 2023; 388:764–6. DOI: 10.1056/NEJMc2215471

92. Liu F, Enanoria WT, Zipprich J, Blumberg S, Harriman K, Ackley SF, et al. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for California. BMC public health 2015; 15:447

93. Long QX, Tang XJ, Shi QL, Li Q, Deng HJ, Yuan J, et al. Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. Nature medicine 2020; 26:1200–4

94. Loomba S, Figueiredo A de, Piatek SJ, Graaf K de, and Larson HJ. Measuring the impact of COVID-19 vaccine misinformation on vaccination intent in the UK and USA. Nature human behaviour 2021; 5:337–48

95. Lothan R, Gutman N, and Yamin D. Country versus pharmaceutical company interests for Hepatitis C treatment. Health Care Management Science 2022; 25:725–49. DOI: 10.1007/s10729-022-09607-2

96. Lucas C, Vogels CBF, Yildirim I, Rothman JE, Lu P, Monteiro V, et al. Impact of circulating SARS-CoV-2 variants on mRNA vaccine-induced immunity. Nature 2021; 600:523–9. DOI: 10.1038/s41586-021-04085-y

97. Lucas M, Karrer U, Lucas A, and Klenerman P. Viral escape mechanisms–escapology taught by viruses. Int J Exp Pathol 2001; 82:269–86. DOI: 10.1046/j.1365-2613.2001.00204.x

98. Mandavilli A. Immunity to the Coronavirus May Last Years, New Data Hint. [Online; accessed 25-November-2020]. 2020 Nov. Available from: https://www.nytimes.com/2020/11/17/health/coronavirus-immunity.html

99. Manley E, Cheng T, Penn A, and Emmonds A. A framework for simulating large-scale complex urban traffic dynamics through hybrid agent-based modelling. Computers, Environment and Urban Systems 2014; 44:27–36

100. Maurice C, Madrigal F, and Lerasle F. Hyper-Optimization Tools Comparison for Parameter Tuning Applications. *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*. IEEE. 2017 :1–6

101. Mazzoni A, Maggi L, Capone M, Vanni A, Spinicci M, Salvati L, et al. Heterogeneous magnitude of immunological memory to SARS-CoV-2 in recovered individuals. Clin Transl Immunology 2021; 10:e1281. DOI: 10.1002/cti2.1281

102. McKay MD, Beckman RJ, and Conover WJ. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics 1979; 21:239–45

103. Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, and Teller E. Equation of State Calculations by Fast Computing Machines. The Journal of Chemical Physics 2004 Dec; 21:1087–92. DOI: 10.1063/1.1699114

104. Mukherjee S and Pahan K. Is COVID-19 Gender-sensitive? Journal of Neuroimmune Pharmacology 2021; 16:38–47. DOI: 10.1007/s11481-020-09974-z

105. Nebehay S. Widespread COVID-19 vaccinations not expected until mid-2021, WHO says. [Online; accessed 25-November-2020]. 2020 Sep. Available from: https://www.reuters.com/article/us-health-coronavirus-who-briefing-idUSKBN25V1B0

106. Ngonghala CN, Iboi E, Eikenberry S, Scotch M, MacIntyre CR, Bonds MH, et al. Mathematical assessment of the impact of non-pharmaceutical interventions on curtailing the 2019 novel Coronavirus. Mathematical Biosciences 2020 :108364

107. Nordstrom P, Ballin M, and Nordstrom A. Risk of SARS-CoV-2 reinfection and COVID-19 hospitalisation in individuals with natural and hybrid immunity: a retrospective, total population cohort study in Sweden. Lancet Infect Dis 2022; 22:781–90. DOI: 10.1016/S1473-3099(22)00143-8

108. Office of Financial Management. Census geographic files - Census 2020. [Online; accessed 24-March-2023]. 2023. Available from: https://ofm.wa.gov/washington-data-research/population-demographics/gis-data/census-geographic-files

109. Oke JB, Akkinepally AP, Chen S, Xie Y, Aboutaleb YM, Azevedo CL, et al. Evaluating the systemic effects of automated mobility-on-demand services via large-scale agent-based simulation of auto-dependent prototype cities. Transportation Research Part A: Policy and Practice 2020; 140:98–126. DOI: https://doi.org/10.1016/j.tra.2020.06.013

110. Panovska-Griffiths J, Kerr CC, Stuart RM, Mistry D, Klein DJ, Viner RM, et al. Determining the optimal strategy for reopening schools, the impact of test and trace interventions, and the risk of occurrence of a second COVID-19 epidemic wave in the UK: a modelling study. The Lancet Child & Adolescent Health 2020; 4:817–27. DOI: 10.1016/S2352-4642(20)30250-9

111. Peng J and Van den Bulte C. Participation vs. Effectiveness in Sponsored Tweet Campaigns: A Quality-Quantity Conundrum. Management Science 2024; 0:null. DOI: 10.1287/mnsc.2019.01897

112. Pew Research Center. Social Media Fact Sheet. [Online; accessed 6-March-2024]. 2024. Available from: https://www.pewresearch.org/internet/fact-sheet/social-media/?tabId=tab-d102dcb7-e8a1-42cd-a04e-ee442f81505a

113. Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, Edmunds J, et al. Dynamic transmission modeling: a report of the ISPOR-SMDM modeling good research practices task force-5. Value in health 2012; 15:828–34

114. Prem K, Cook AR, and Jit M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS computational biology 2017; 13:e1005697

115. Rao IJ and Brandeau ML. Sequential Allocation of Vaccine to Control an Infectious Disease. Mathematical Biosciences 2022; 351:108879. DOI: https://doi.org/10.1016/j.mbs.2022.108879

116. Reeves DB, Bracis C, Swan DA, Burns E, Moore M, Dimitrov D, et al. Rapid vaccination and partial lockdown minimize 4th waves from emerging highly contagious SARS-CoV-2 variants. Med 2021; 2:573–4. DOI: 10.1016/j.medj.2021.04.012

117. Reiner RC, Barber RM, Collins JK, Zheng P, Adolph C, Albright J, et al. Modeling COVID-19 scenarios for the United States. Nature Medicine 2020. DOI: 10.1038/s41591-020-1132-9

118. Roy DN, Biswas M, Islam E, and Azam MS. Potential factors influencing COVID-19 vaccine acceptance and hesitancy: A systematic review. PLOS ONE 2022 Mar; 17:1–20. DOI: 10.1371/journal.pone.0265496

119. SafeGraph. The impact of coronavirus (COVID-19) on foot traffic. [Online; accessed 24-March-2023]. 2020. Available from: https://www.safegraph.com/dashboard/covid19-commerce-patterns

120. Santos JV, Gomes da Costa J, Costa E, Almeida S, Cima J, and Pita-Barros P. Factors associated with non-pharmaceutical interventions compliance during COVID-19 pandemic: a Portuguese cross-sectional survey. J Public Health (Oxf) 2023; 45:47–56. DOI: 10.1093/pubmed/fdac001

121. Sanz-Leon P, Hamilton LHW, Raison SJ, Pan AJX, Stevenson NJ, Stuart RM, et al. Modelling herd immunity requirements in Queensland: impact of vaccination effectiveness, hesitancy and variants of SARS-CoV-2. Philos Trans A Math Phys Eng Sci 2022; 380:20210311. DOI: 10.1098/rsta.2021.0311

122. Sheehan MM, Reddy AJ, and Rothberg MB. Reinfection rates among patients who previously tested positive for coronavirus disease 2019: a retrospective cohort study. Clinical Infectious Diseases 2021; 73:1882–6

123. Stein C, Nassereldine H, Sorensen RJD, Amlag JO, Bisignano C, Byrne S, et al. Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis. The Lancet 2023; 401:833–42. DOI: 10.1016/S0140-6736(22)02465-5

124. Steinert JI, Sternberg H, Prince H, Fasolo B, Galizzi MM, Büthe T, et al. COVID-19 vaccine hesitancy in eight European countries: Prevalence, determinants, and heterogeneity. Science Advances 2022; 8:eabm9825. DOI: 10.1126/sciadv.abm9825

125. Tatapudi H, Das R, and Das TK. Impact assessment of full and partial stay-at-home orders, face mask usage, and contact tracing: An agent-based simulation study of COVID-19 for an urban region. Glob Epidemiol 2020; 2:100036. DOI: 10.1016/j.gloepi.2020.100036

126. Thakkar N, Burstein R, Klein D, and Famulare M. Sustained reductions in transmission have led to declining COVID- 19 prevalence in King County, WA. 2020 Apr. Available from: https://covid.idmod.org/data/Sustained_reductions_in_transmission_have_led_to_declining_COVID_19_prevalence_in_King_County_WA.pdf

127. Thompson RN, Hill EM, and Gog JR. SARS-CoV-2 incidence and vaccine escape. Lancet Infect Dis 2021; 21:913–4. DOI: 10.1016/S1473-3099(21)00202-4

128. Thul L and Powell W. Stochastic optimization for vaccine and testing kit allocation for the COVID-19 pandemic. European Journal of Operational Research 2023; 304. The role of Operational Research in future epidemics/ pandemics:325–38. DOI: https://doi.org/10.1016/j.ejor.2021.11.007

129. Tseng HF, Ackerson BK, Luo Y, Sy LS, Talarico CA, Tian Y, et al. Effectiveness of mRNA-1273 against SARS-CoV-2 Omicron and Delta variants. Nat Med 2022; 28:1063–71

130. Tučník P and Bureš V. Experimental Evaluation of Suitability of Selected Multi-Criteria Decision-Making Methods for Large-Scale Agent-Based Simulations. PLOS ONE 2016 Nov; 11:1–24. DOI: 10.1371/journal.pone.0165171

131. UK Health Security Agency. COVID-19 vaccine surveillance report - Week 11. Web Page. [Online; accessed 14-June-2022]. 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061532/Vaccine_surveillance_report_-_week_11.pdf

132. UK Health Security Agency. COVID-19 vaccine surveillance report - Week 24. Web Page. [Online; accessed 14-September-2022]. 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1083443/Vaccine-surveillance-report-week-24.pdf

133. Ulloa AC, Buchan SA, Daneman N, and Brown KA. Estimates of SARS-CoV-2 Omicron Variant Severity in Ontario, Canada. JAMA 2022; 327:1286–8. DOI: 10.1001/jama.2022.2274

134. Vanni T, Karnon J, Madan J, White RG, Edmunds WJ, Foss AM, et al. Calibrating Models in Economic Evaluation. PharmacoEconomics 2011; 29:35–49. DOI: 10.2165/11584600-000000000-00000

135. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases 2020. DOI: 10.1016/S1473-3099(20)30243-7

136. Viana J, Dorp CH van, Nunes A, Gomes MC, Boven M van, Kretzschmar ME, et al. Controlling the pandemic during the SARS-CoV-2 vaccination rollout. Nat Commun 2021; 12:3674. DOI: 10.1038/s41467-021-23938-8

137. Vilches TN, Sah P, Moghadas SM, Shoukat A, Fitzpatrick MC, Hotez PJ, et al. COVID-19 hospitalizations and deaths averted under an accelerated vaccination program in northeastern and southern regions of the USA. Lancet Reg Health Am 2022; 6:100147. DOI: 10.1016/j.lana.2021.100147

138. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 2020; 17:261–72. DOI: 10.1038/s41592-019-0686-2

139. Walls AC, Sprouse KR, Bowen JE, Joshi A, Franko N, Navarro MJ, et al. SARS-CoV-2 breakthrough infections elicit potent, broad, and durable neutralizing antibody responses. Cell 2022; 185:872–880 e3. DOI: 10.1016/j.cell.2022.01.011

140. Washington State, Department of Health. COVID-19 data dashboard. [Online; accessed 24-March-2023]. 2023. Available from: https://doh.wa.gov/emergencies/covid-19/data-dashboard

141. Washington State, Department of Health. Roadmap to Recovery Metrics. Web Page. [Online; accessed 20-June-2021]. 2021. Available from: https://coronavirus.wa.gov/what-you-need-know/roadmap-recovery-metrics

142. Washington State, Department of Health. SARS-CoV-2 Sequencing and Variants in Washington State. Web Page. [Online; accessed 6-September-2022]. 2022. Available from: https://doh.wa.gov/sites/default/files/2022-02/420-316-SequencingAndVariantsReport.pdf?uid=63179ee18bd2e

143. Washington State, Department of Health. Summary of COVID-19 vaccination among King County residents. Web Page. [Online; accessed 6-July-2022]. 2022. Available from: https://kingcounty.gov/depts/health/covid-19/data/vaccination.aspx

144. Washington State, Department of Health. Washington's COVID-19 Vaccine Phases. Web Page. [Online; accessed 1-June-2021]. 2021. Available from: https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/VaccinationPhasesInfographic.pdf

145. Washington State, The Office of The Governor. Safe Start Washington – Phased Reopening County-by-County. [Online; accessed 22-October-2020]. 2020. Available from: https://www.governor.wa.gov/sites/default/files/SafeStartPhasedReopening.pdf

146. Wheaton W and RTI International. US Synthetic Population 2010 Version 1.0: Quick Start Guide. [Online; accessed 1-December-2020]. 2014. Available from: https://www.rti.org/impact/rti-us-synthetic-household-population%E2%84%A2

147. Wong MC, Wong EL, Huang J, Cheung AW, Law K, Chong MK, et al. Acceptance of the COVID-19 vaccine based on the health belief model: A population-based survey in Hong Kong. Vaccine 2021; 39:1148–56

148. World Health Organization. Interim statement on hybrid immunity and increasing population seroprevalence rates. Web Page. [Online; accessed 5-October-2022]. 2022. Available from: https://www.who.int/news/item/01-06-2022-interim-statement-on-hybrid-immunity-and-increasing-population-seroprevalence-rates

149. Xie W, Zhang P, and Zhang Q. A Stochastic Simulation Calibration Framework for Real-Time System Control. *Proceedings of the 2017 Winter Simulation Conference*. Ed. by Chan WKV, D'Ambrogio A, Zacharewicz G, Mustafee N, Wainer G, and Page EH. Institute of Electrical and Electronics Engineers, Inc. Las Vegas, NV, USA, 2017 :1914–25. DOI: 10.1109/WSC.2017.8247927

150. Yamin D and Gavious A. Incentives' Effect in Influenza Vaccination Policy. Management Science 2013; 59:2667–86. DOI: 10.1287/mnsc.2013.1725

151. Yin Q, Wang Z, Xia C, and Bauch CT. Impact of co-evolution of negative vaccine-related information, vaccination behavior and epidemic spreading in multilayer networks. Communications in Nonlinear Science and Numerical Simulation 2022; 109:106312

152. Young BE, Ong SWX, Kalimuddin S, Low JG, Tan SY, Loh J, et al. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. JAMA 2020; 323:1488–94

153. Zhang J and Centola D. Social Networks and Health: New Developments in Diffusion, Online and Offline. Annual Review of Sociology 2019; 45:91–109. DOI: 10.1146/annurev-soc-073117-041421

154. Zhou Y, Xu R, Hu D, Yue Y, Li Q, and Xia J. Effects of human mobility restrictions on the spread of COVID-19 in Shenzhen, China: a modelling study using mobile phone data. Lancet Digit Health 2020; 2:e417–e424. DOI: 10.1016/S2589-7500(20)30165-5

Appendix A

APPENDIX OF CHAPTER 2

Contents

A.1 Appendix of Section 2.2, Impact of nonpharmaceutical interventions on reopening society

A.1.1 COVID-19 Model Specification and Data Input under Original Wuhan Strain

Disease Parameters

We estimated COVID-19 disease parameters from published literature and reported data. As shown in Table A.1.1, we considered values of R_0 between 1.5 to 3.5 [5] and calibrated transmissibility of SARS-CoV-2. We used a lognormal distribution with median of 5.1 days [28] for the incubation period. We set the infectious period to 15 days [20, 53]. For symptomatic individuals, the infectious period consists of 2 presymptomatic days [20] and 13 symptomatic days [43]. See Figure A.1.1 for details. Asymptomatic individuals have the same infectious period of 15 days as symptomatic individuals. As suggested in the literature [3, 20, 53], we used a sigmoid curve to model the daily infectivity that peaks on first symptom onset day and then steadily decreases. Also, asymptomatic individuals' infectiousness is assumed to be 75% that of symptomatic individuals [5].

The probability of an individual being asymptomatic is by age group [11], and the probability of severe symptoms is by age group [5]. The probability that an infectious individual dies (infection fatality ratio) is based on [5] and the probability that a symptomatic individual dies (infection fatality ratio among symptomatic) is based on [4].

	Value	Reference
Disease Parameters		
Reproduction number (R_0)	1.5-3.5	[5]
Transmissibility of SARS-CoV-2	5.1% (Cluster 2), 8.9% (Cluster 6)	Calibrated
Incubation period distribution	Lognormal with median of 5.1 days, dispersion of 1.52 days	[28]
Symptomatic period	13 days	[43]
Pre-symptomatic period	2 days	[20]
Infectious period	15 days	[20, 53]
Daily infectivity	0.5, 0.69, 0.94, 0.8, 0.61, 0.47, 0.35, 0.27, 0.2, 0.16, 0.12, 0.09, 0.07, 0.05,0.04	[3, 20, 53]
Infectiousness of asymptomatic individuals relative to symptomatic	0.75	[5]
Age group for asymptomatic infection†	0-9, 10-19, 20-29, 30-39, 40-49, 50-59, 60-69, 70+	[11]
Probability of asymptomatic infection by age group†	0.71, 0.79, 0.73, 0.67, 0.6, 0.51, 0.37, 0.31	[11]
Probability of severe symptoms by age group	0-49:0.017, 50-64:0.045, 65+:0.074	[5]
Probability of acquiring immunity after recovered	1	Assumed
Days from symptom onset to death	15 days	[5]
Infection fatality ratio	0.005-0.008 (Best estimate: 0.0065)	[5]
Infection fatality ratio among symptomatic	0.005-0.061	[4]

† For example, 10 year-old individuals have 79% chance of becoming symptomatic.

Table A.1.1: COVID-19 Disease Parameters.

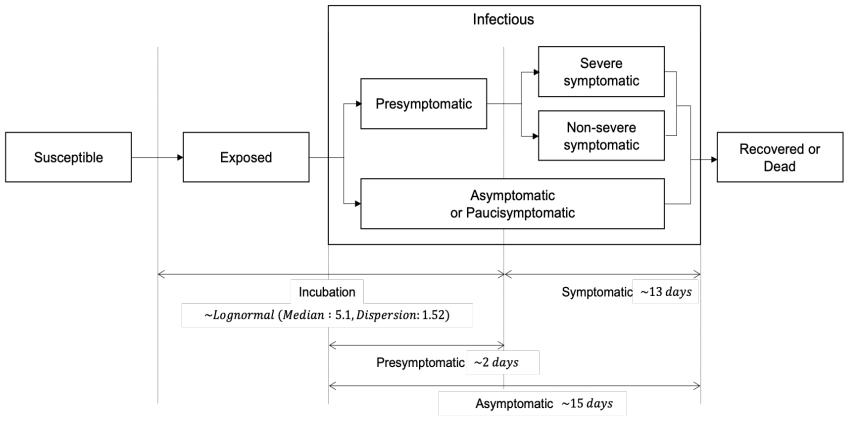


Figure A.1.1: Natural history of SARS-CoV-2 infection.

Parameter Calibration

We calibrated six unknown model parameters, which are transmissibility of SARS-CoV-2, daily contact probability between household members, daily contact rates of neighborhood, school, and workplace, and default home quarantine percentage of symptomatic individuals. In the first step, we used the Latin hypercube sampling method to sample 1,000 candidate parameter sets. We initially infected 10 random persons and simulated for 30 days, which is a period long enough for the initial infectors to complete their incubation and infectious period. We then calculated the average number of secondary infections caused by the initial infectors, that is, R_0. This resulted in selecting 65 parameter sets that achieved a R_0 between 1.5 and 3.5 [5].

Using Minitab 19.2020.1 (Minitab LLC), we performed complete-linkage clustering of the 65 selected parameter sets to identify clusters. The complete-linkage clustering is one of the methods of hierarchical clustering. The method starts with 65 clusters, and iteratively combines two clusters that have the farthest pair among all pairs of parameter sets as it decreases the number of clusters. It is recommended to stop when combining another pair of clusters results in an abrupt decrease of similarity level within each cluster [32]. In our model, this resulted in identifying 7 clusters.

In the second step of our calibration process, we simulated public health interventions in King County from January 15, 2020 through May 31, 2020 following Washington State government's "Stay Home, Stay Healthy" emergency order (Figure A.1.6). We then selected clusters that met both the reported Infection Fatality Ratio (IFR) and reported Infection Fatality Ratio among symptomatic cases (IFR-S) using ranges found in the literature [4, 5] (Table A.1.1). IFR was calculated as the ratio between the reported deaths and simulated total infections, while the IFR-S was calculated as the ratio between the reported deaths and simulated symptomatic infections. This resulted in selecting clusters 2 and 6. Although we only targeted cumulative deaths until May 31, 2020, Figure A.1.2 shows that the trends in daily projected deaths using IFR are similar to daily reported deaths until May 31, 2020. Cluster 2 represents higher contact rates in household, neighborhood, and workplace. Cluster 6 represents higher overall transmissibility of SARS-CoV-2, school contact rates, and default home quarantine percentage when symptomatic (Table A.1.2).

Cluster	Calibrated parameters						First step calibration result	Second step calibration result	
	Transmissibility	Daily household contact probability	Daily contact rate in neighborhood	Daily contact rate in school	Daily contact rate in workplace	Home quarantine percentage when symptomatic	R_0	Total Infections until May 31th	Symptomatic infections until May 31th
1	7.2%	33%	4.71	3.34	1.56	64%	2.50	7,218	2,570
2	5.1%	81%	4.18	4.00	5.08	57%	2.40	113,989	40,905
3	16.9%	58%	1.53	4.06	1.20	78%	3.10	667,975	236,648
4	5.4%	36%	5.59	7.20	2.77	45%	2.45	56,613	20,374
5	9.0%	58%	1.75	4.51	1.28	26%	2.20	16,393	5,556
6	8.9%	52%	2.39	6.64	1.82	78%	2.40	82,073	28,173
7	7.8%	22%	2.19	3.56	5.19	53%	2.60	158,032	59,902
Calibration ranges	1-100%	10-100%	0.5-7	1.5-9	0-7	0-100%			
References	Assumed	[31]	[31]	[31]	[31]	[31]			

Table A.1.2: Calibration Result.

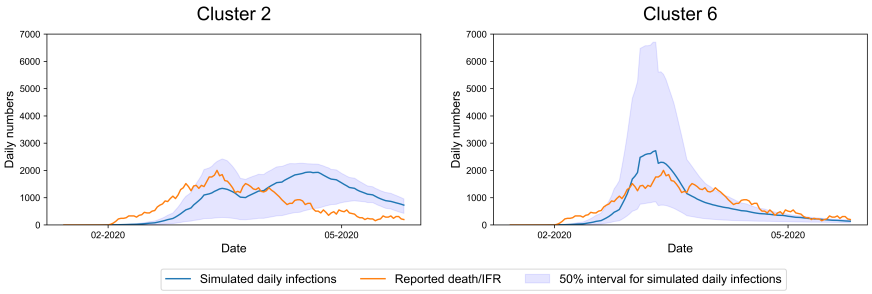


Figure A.1.2: Calibration to reported deaths. 50% confidence interval was obtained from calculating 25%th and 75%th percentile at each time step. IFR: Infection Fatality Ratio

Intervention Parameters

	Base Case Value	Reference
Intervention Parameters		
Social distancing level	20%	Assumed
Total diagnosed cases per 100K population in two weeks to enter phase 2,3,4[†]	Under 50, 25, 10	[50]
Face mask compliance	50%	Assumed
Effectiveness of wearing face mask	50%	[13]
School closure	50%	Assumed
Home quarantine percentage when symptomatic[#]	57% (Cluster 2), 78% (Cluster 6)	Calibrated
Home quarantine duration when contact traced	14 days since last exposure	[7]
Percent tested of severe symptomatic individuals[§]	100%	Assumed
Percent tested of non-severe symptomatic individuals[§]	60%	Assumed
Delay from onset to testing	2 days	[34]
Delay to receive test results	1 day	[34]
Test sensitivity	0.9	Assumed
Infector's compliance with contact tracing	0.9	[50]
Contact tracing effectiveness (Household, school, workplace, neighborhood member's compliance with home quarantine and testing when contact traced)[§]	0.9, 0.72, 0.16, 0.10	Assumed
Delays in reaching contacts	2 days	[50]

[†] Applied for phase-based social distancing. If the value is over 50, phase 1 is applied. We assumed that phase 1,2,3, and 4 corresponds to high, medium, low, and no social distancing.
[#] We applied the same probability of staying home when symptomatic to severe and non-severe symptomatic individuals. Because symptom-severity gradually increases, severe symptomatic individuals may have mild symptoms initially but suffer longer and have more severe symptoms later on
[§] See text for calculation details.

Table A.1.3: Intervention Parameters.

To calculate percent tested for severe and non-severe symptomatic individuals, we defined testing effectiveness based on the testing percentage of non-severe symptomatic individuals; low (30%), medium (60%), and high (90%). The baseline value for the testing percentage (medium, 60%) was calculated based on King County's testing history. We assumed severe symptomatic individuals were all tested since March 1, 2020, and some percentage of non-severe symptomatic individuals were tested since March 20, 2020. To calculate the percentage, we compared case fatality rate with estimated infection fatality ratio among symptomatic cases until May 31, 2020 in King County. We assumed that King County only tested symptomatic individuals and all deaths

caused by COVID-19 were reported. In this case,

$$\text{Case fatality rate} = \frac{\text{Reported death}}{\text{Reported cases}} = \frac{\text{Deceased individuals}}{(\text{Symptomatic individuals}) * P(\text{test for symptomatic individuals})}$$
$$= \frac{\text{IFR-S}}{P(\text{test for symptomatic individuals})}.$$

Then, $P(\text{test for symptomatic individuals}) = \text{IFR-S}/\text{Case fatality rate}$. Using this equation, we calculated that 65% of symptomatic individuals tested positive. As we assumed around 3.4% of symptomatic individuals are severe [5], we concluded that around 60% of non-severe symptomatic individuals are tested.

As seen from Figure A.1.3, the effects of increasing contact tracing compliance are marginal compared to other NPIs such as social distancing, face mask use, and school closure. We did not observe clear red or green combinations that result in a surge of second wave or elimination of the pandemic. Therefore, for simplicity, we combined these two interventions into one intervention with three levels; low testing and low contact tracing, medium testing and medium contact tracing, and high testing and high contact tracing.

Contact tracing effectiveness was defined as the compliance rate for home quarantine and testing when contact traced. We calculated this based on [27] through multiplying the probability that contacts are acquaintances and the probability that the contacts are traceable for each place that the contact occurred. Contact tracing effectiveness was categorized as low, medium, and high where low effectiveness means tracing only household members, medium level means tracing most household and school members but limited tracing to workplace and neighborhood members, and high effectiveness means tracing most acquaintances of household, schools, workplace, and neighborhood (Table A.1.4).

High effective contact tracing	P(Acquiantances)	P(Traceable)	Home quarantine and testing compliance when contact traced
Household	100%	100%	100%
School	90%	95%	86%
Workplace	79%	95%	75%
Neighborhood	52%	95%	49%

Medium effective contact tracing	P(Acquiantances)	P(Traceable)	Home quarantine and testing compliance when contact traced
Household	100%	90%	90%
School	90%	80%	72%
Workplace	79%	20%	16%
Neighborhood	52%	20%	10%

Low effective contact tracing	P(Acquiantances)	P(Traceable)	Home quarantine and testing compliance when contact traced
Household	100%	90%	90%
School	90%	0%	0%
Workplace	79%	0%	0%
Neighborhood	52%	0%	0%

Table A.1.4: Values used to calculate the contact tracing effectiveness equation

Figure A.1.3: Effects of different combinations of testing and contact tracing on total infections (1,000s) from June 2020 to May 2021

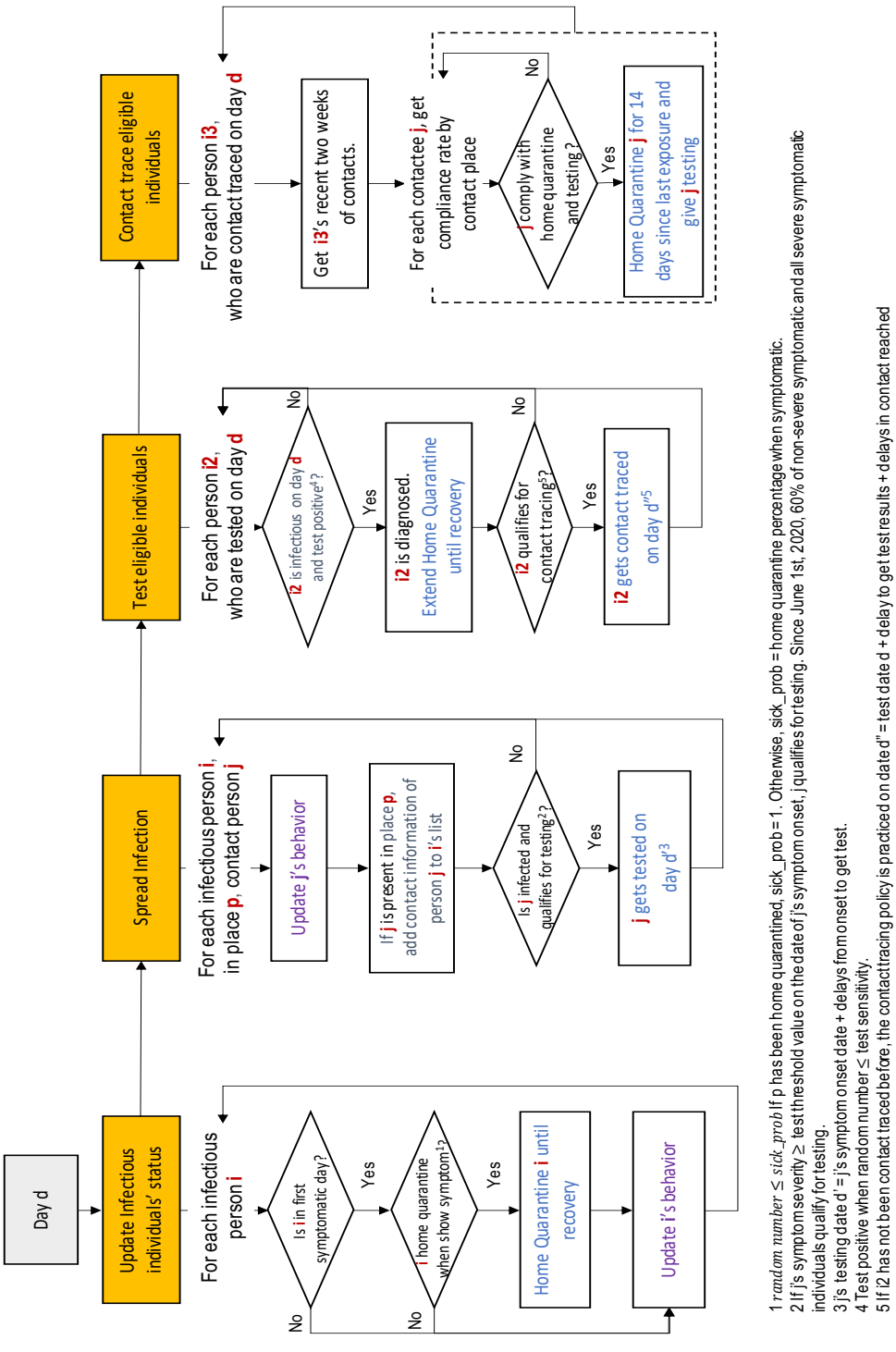
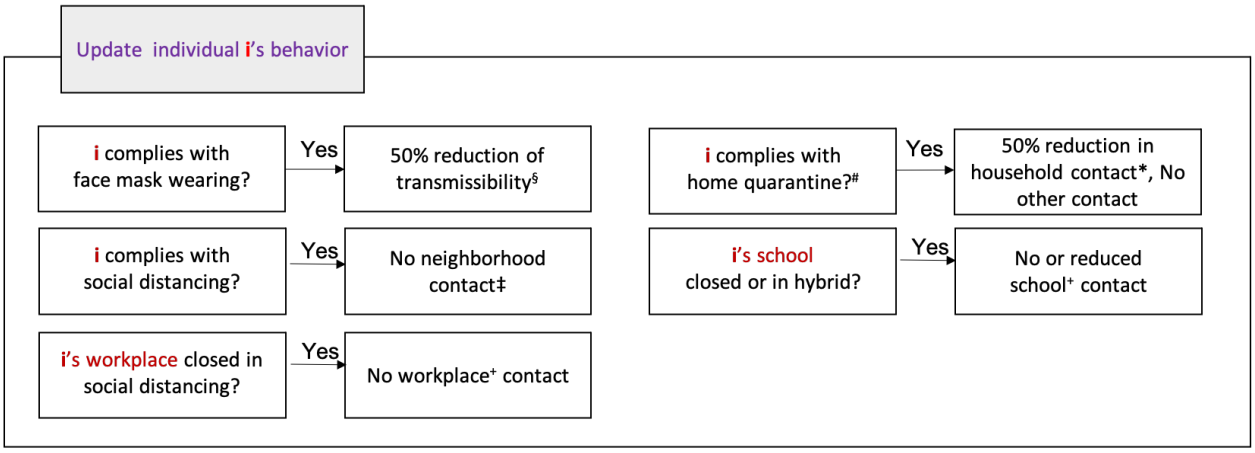


Figure A.1.4: Daily flowchart for viral transmission and interventions. Blue texts identify public health interventions. Purple texts indicate individuals' daily behavioral decisions that are impacted by NPIs, and are expanded in Figure A.1.5.



§ Not applied in household contact. + Workplace contacts include office contact which is a smaller mixing network with twice of the workplace contact rate per day. Same policy is applied for school which includes classroom contact. ‡ Both infectious and contact person may not have neighborhood contact, which can have double-effect on reducing viral transmission. # Do home quarantine when symptomatic individuals feel sick (by default), get diagnosed, or contact traced. * Before March 1, 2020, there is no reduction in household contact.

Figure A.1.5: Effects of NPIs on individuals' behavior.

SARS-CoV-2 Transmission Equations

For an infectious person i, we consider his or her household transmission for each dth infectious day. For each susceptible household member j, j becomes infected with Household infection probability (i,j,d):

$$\begin{aligned}\text{Household infection probability } (i,j,d) = (&\text{daily household contact probability}) \times (1 - \text{reduction} \\ &\text{in contact if } i \text{ is home quarantined}) \\ &\times (\text{transmissibility of SARS-CoV-2}) \\ &\times \text{daily infectivity}(i,d) \times (\text{reduced infectiousness} \\ &\text{if } i \text{ is asymptomatic}).\end{aligned}$$

In a neighborhood, a workplace, or a school, if an infectious person i is present at the place p on dth infectious day, the expected number of contacts is the Contact count (i, p, d):

Contact count (i, p, d) = (daily contacts per day (p)) × (1 - reduction in contact if in partially-open school) × (transmissibility of SARS-CoV-2) × daily infectivity (i, d) × (reduced infectiousness if i is asymptomatic).

If Contact count (i,p,d) is not integer, stochastic rounding is needed. If Contact count (i,p,d) $= x$, then

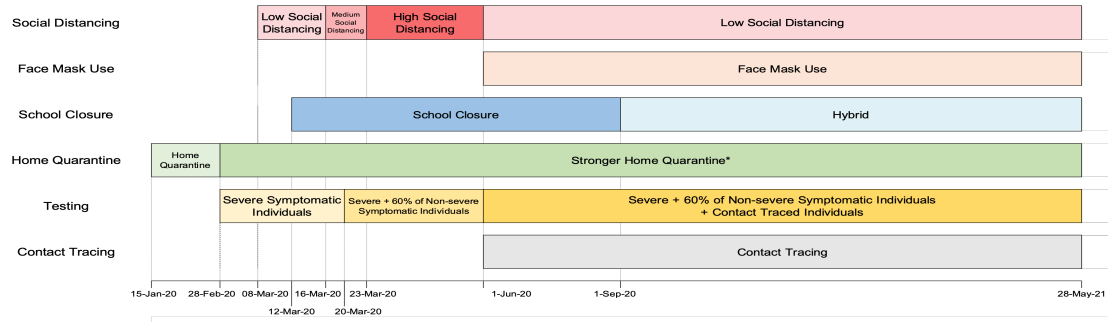
$$\text{Actual contact count}(i,p,d) = Round(x) = \begin{cases} \lfloor x \rfloor \text{ with probability } 1-(x-\lfloor x \rfloor) \\ \lfloor x \rfloor + 1 \text{ with probability } (x-\lfloor x \rfloor) \end{cases}$$

For example, if $x = 2.7$, Actual contact count(i,p,d) would be 2 with probability 0.3 and 3 with probability 0.7. We randomly select Actual contact count (i,p,d) individuals in place p on day d. For each contactee j in the place p, if j is susceptible and present in the place, j becomes infected with Non-household infection probability (i,j):

Non-household infection probability (i,j) = (1 - effectiveness of wearing face mask if i uses mask)× (1 - effectiveness of wearing face mask if j uses mask).

A.1.2 Base Case Timeline

We assumed the first infection occurred on January 15, 2020 by a man who lived in Snohomish County and commuted to Seattle, and thus impacted King County. We set the initial infector's age and isolation date to this man's profile [21]. We also made a conservative assumption in the base case that this is the only non-local transmission and all other transmissions are community transmissions in King County.



* Stronger home Quarantine: 50% reduction in household contact.

Figure A.1.6: Diagram of Simulated Interventions for Base Case.

A.1.3 Scenario Objective Calculated from Washington State's Objective

We translated Washington State's objective of having less than 25 diagnosed cases per 100,000 people per 14 days into total infections from June 1, 2020 to May 28, 2021 (362 days) using Equation (A.1). The equation consists of King County's population in the model (1,893,508), maximum total diagnosed cases (25) in 14-day period, and the assumption that 40% of infections that are diagnosed as estimated in the base case result.

$$\text{Objective in total infections from June 2020 to May 2021} = \frac{1,893,508}{100,000} \times \frac{25 \times 362}{14} \div 0.4 \approx 30,600 \tag{A.1}$$

A.1.4 One-way Scenario Analysis Result

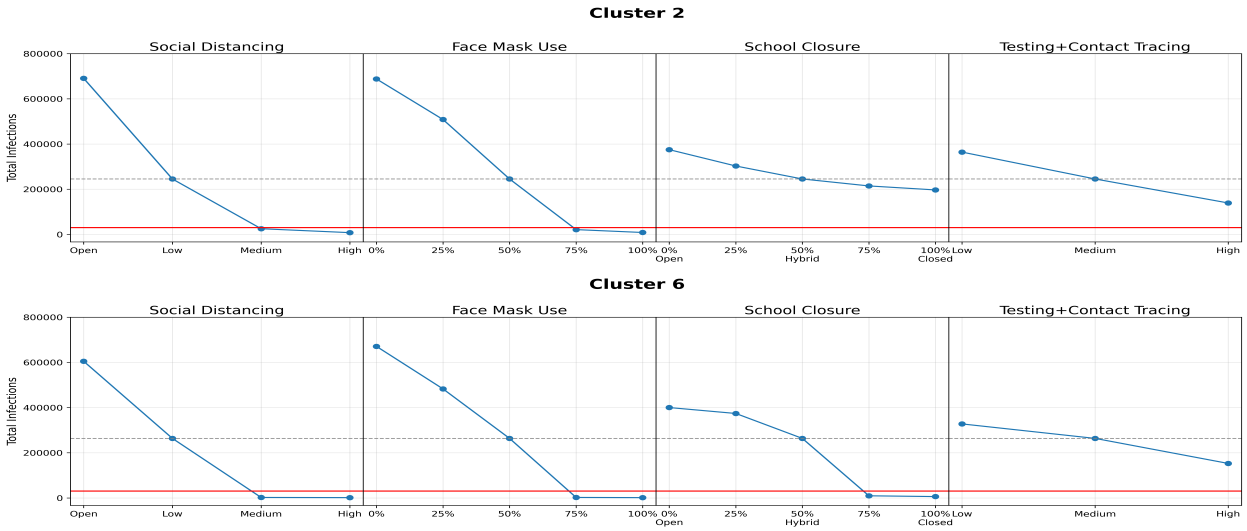


Figure A.1.7: Results of one-way scenario analyses from NPIs only. Total infections from June 1, 2020 to May 28, 2021. Red lines indicate Washington State's objective value and dotted lines indicate the base case result.

Cluster 2	Value	Total Infection June 1 2020 to May 28 2021	Peak Incidence	Peak Incidence date	Quarantine percentage for infectious person	Average daily test	Average daily contact trace	Average daily quarantine
Base case		245,448	1,428	2020/09/29	33%	717	233	7,969
Base case - 50% Interval		[135,576, 336,192]	[847, 1,899]	[2020-08-13, 2020-10-20]	[33%,33%]	[398, 1006]	[130, 328]	[4,465, 11,095]
Daily imported infections since June 1 2020	**0**	245,448	1,428	2020/09/29	33%	717	233	7,969
	1	259,810	1,450	2020/09/26	33%	760	248	8,382
	10	283,009	1,610	2020/10/02	33%	822	268	9,057
Intervention parameters								
Social Distancing	Open	690,835	10,805	2020/07/25	42%	3,110	1,205	28,980
	Low	245,448	1,428	2020/09/29	33%	717	233	7,969
	Medium	25,289	728	2020/05/26	32%	85	27	1,013
	High	7,922	713	2020/05/26	31%	134	48	1,310
School Closure	0%	375,206	2,653	2020/10/21	32%	1,102	351	11,948
	25%	302,879	1,844	2020/10/15	32%	886	285	9,742
	50%	245,448	1,428	2020/09/29	33%	717	233	7,969
	75%	214,673	1,291	2020/09/18	33%	626	206	7,034
	100%	196,940	1,228	2020/09/12	33%	574	190	6,477
Face Mask Use	0%	688,032	13,220	2020/07/22	34%	2,317	883	22,781
	25%	508,587	6,206	2020/08/08	33%	1,465	547	15,496
	50%	245,448	1,428	2020/09/29	33%	717	233	7,969
	75%	21,336	754	2020/05/27	32%	103	34	1,033
	100%	8,638	728	2020/05/26	31%	157	51	1,415
Testing+Contact Tracing	Low	364,402	3,062	2020/09/09	22%	366	130	5,894
	Medium	245,448	1,428	2020/09/29	33%	717	233	7,969
	High	139,224	744	2020/05/27	40%	607	196	6,318
Face Mask Effectiveness	25%	505,481	5,664	2020/08/10	33%	1,500	537	15,481
	50%	245,448	1,428	2020/09/29	33%	717	233	7,969
	75%	47,838	754	2020/05/27	32%	153	48	1,762
	100%	16,777	727	2020/05/27	32%	84	28	824
Delay from onset to testing	1 day	222,119	1,195	2020/10/04	34%	659	215	7,425
	2 days	245,448	1,428	2020/09/29	33%	717	233	7,969
	5 days	296,385	2,155	2020/09/16	31%	822	260	8,615
Delay to receive test results	**1 day**	245,448	1,428	2020/09/29	33%	717	233	7,969
	3 day	282,796	1,929	2020/09/26	31%	756	241	8,321
	5 days	312,818	2,313	2020/09/20	29%	761	244	8,379
Test sensitivity	0.8	264,370	1,707	2020/09/27	31%	688	211	7,598
	0.9	245,448	1,428	2020/09/29	33%	717	233	7,969
	1	218,061	1,171	2020/09/30	36%	716	245	7,951
Delays in reaching contacts	1 day	220,937	1,183	2020/09/27	35%	676	217	7,555
	2 days	245,448	1,428	2020/09/29	33%	717	233	7,969
	5 days	298,560	2,149	2020/09/16	30%	770	243	8,405
Infector's compliance with contact tracing	0.6	295,842	2,046	2020/09/17	29%	580	153	6,975
	0.8	263,356	1,624	2020/09/19	32%	677	208	7,703
	0.9	245,448	1,428	2020/09/29	33%	717	233	7,969

Table A.1.5: Full results of one-way scenario analyses (Cluster 2). Base case in bold.

Cluster 6	Value	Total Infection June 1 2020 to May 28 2021	Peak Incidence	Peak Incidence date	Quarantine percentage for infectious person	Average daily test	Average daily contact trace	Average daily quarantine
Base case		263,595	2,979	2021/01/20	34%	767	238	8,506
Base case - 50% Interval		[88,061, 482,395]	[1326, 4,703]	[2020-12-31, 2021-02-17]	[34%,34%]	[254, 1428]	[79, 442]	[2,842, 15,623]
Daily imported infections since June 1 2020	**0**	263,595	2,979	2021/01/20	34%	767	238	8,506
	1	291,331	3,008	2021/01/20	34%	847	262	9,335
	10	330,594	3,256	2020/12/25	34%	969	300	10,637
Intervention parameters								
Social Distancing	Open	604,945	10,386	2020/10/11	42%	2,380	891	23,830
	Low	263,595	2,979	2021/01/20	34%	767	238	8,506
	Medium	2,197	134	2020/05/26	34%	35	12	348
	High	1,497	130	2020/05/26	34%	37	13	382
School Closure	0%	400,149	8,840	2020/11/11	36%	1,396	454	13,754
	25%	373,913	6,239	2020/11/28	34%	1,131	359	12,331
	50%	263,595	2,979	2021/01/20	34%	767	238	8,506
	75%	9,584	142	2020/05/27	34%	28	9	335
	100%	6,143	142	2020/05/27	34%	31	11	346
Face Mask Use	0%	670,793	14,371	2020/09/30	35%	2,235	830	22,229
	25%	482,661	8,307	2020/10/31	35%	1,376	477	15,030
	50%	263,595	2,979	2021/01/20	34%	767	238	8,506
	75%	2,267	137	2020/05/27	34%	33	11	343
	100%	1,443	133	2020/05/26	34%	38	14	398
Testing+Contact Tracing	Low	327,689	4,693	2020/12/17	24%	333	113	5,648
	Medium	263,595	2,979	2021/01/20	34%	767	238	8,506
	High	152,786	1,712	2021/03/16	39%	642	195	6,398
Face Mask Effectiveness	25%	467,097	7,809	2020/11/07	35%	1,312	458	14,568
	50%	263,595	2,979	2021/01/20	34%	767	238	8,506
	75%	3,185	137	2020/05/27	34%	11	4	144
	100%	1,762	133	2020/05/26	34%	33	11	324
Delay from onset to testing	1 day	214,968	2,449	2021/02/10	35%	646	199	7,177
	2 days	263,595	2,979	2021/01/20	34%	767	238	8,506
	5 days	285,406	3,607	2021/01/03	32%	757	229	8,043
Delay to receive test results	**1 day**	263,595	2,979	2021/01/20	34%	767	238	8,506
	3 day	294,688	3,686	2021/01/06	32%	763	233	8,445
	5 days	310,965	4,079	2020/12/24	30%	724	222	8,046
Test sensitivity	0.8	277,350	3,23	2021/01/13	32%	713	209	7,975
	0.9	263,595	2,979	2021/01/20	34%	767	238	8,506
	1	248,272	2,479	2021/01/28	36%	830	267	9,022
Delays in reaching contacts	1 day	230,250	2,465	2021/01/28	35%	712	215	7,867
	2 days	263,595	2,979	2021/01/20	34%	767	238	8,506
	5 days	291,054	3,829	2020/12/30	31%	724	219	7,949
Infector's compliance with contact tracing	0.6	270,116	3,572	2021/01/07	30%	512	130	6,274
	0.8	241,971	3,192	2021/01/20	33%	621	182	7,011
	0.9	263,595	2,979	2021/01/20	34%	767	238	8,506

Table A.1.6: Full results of one-way scenario analyses (Cluster 6). Base case in bold.

A.1.5 Two-way Scenario Analysis on Testing and Contact Tracing

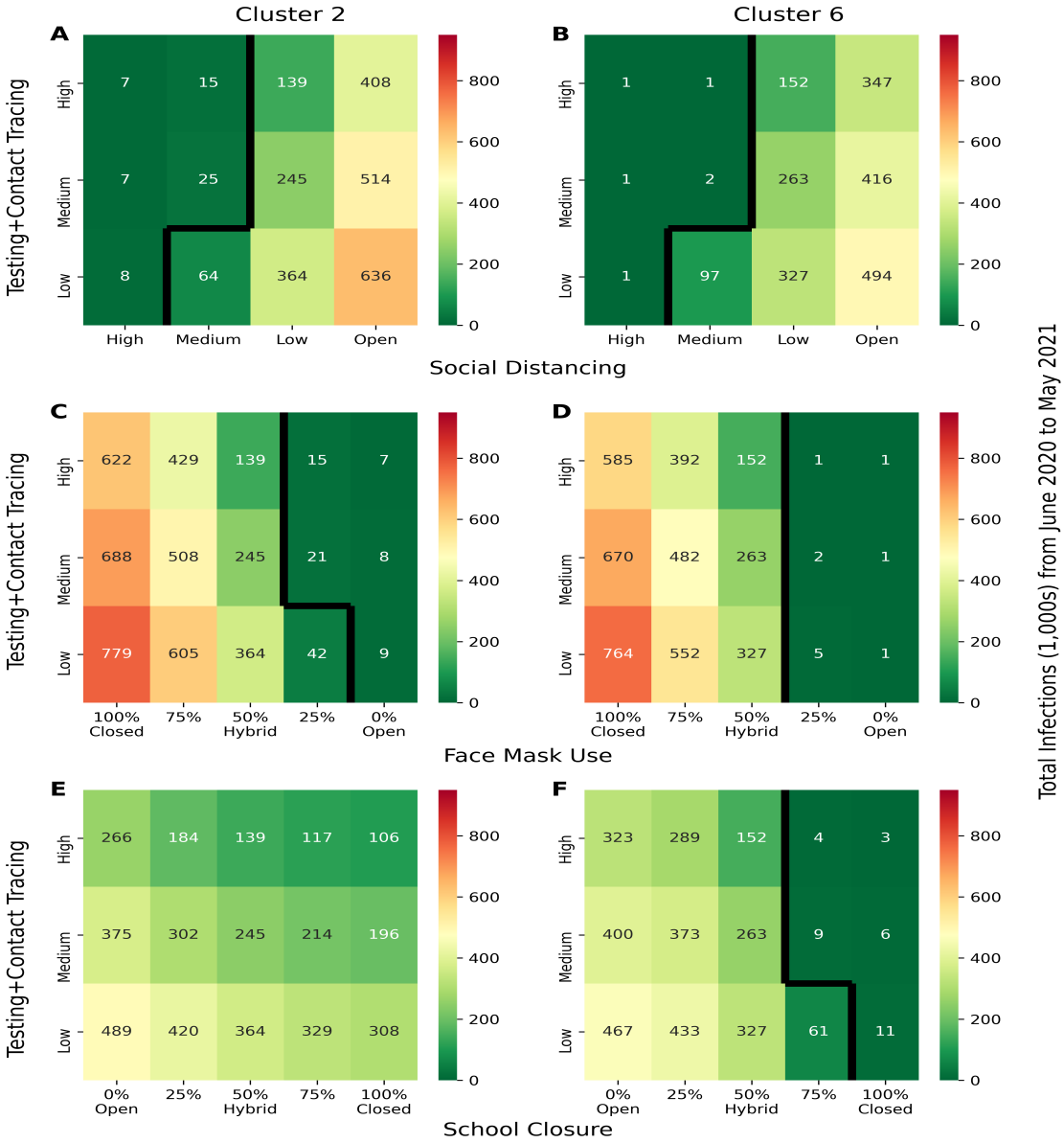


Figure A.1.8: Two-way scenario analysis between testing+contact tracing and other NPIs (social distancing, face mask use, and school closure. Numbers represent total infections (1,000s) from June 1, 2020 to May 28, 2021. Black lines indicate combinations that first satisfy Washington State's objective of 30.6 thousand total infections. In (E), any combination of testing and contact tracing and school closure cannot achieve the objective without stricter social distancing and/or face mask use.

A.1.6 Four-way Scenario Analysis Result

We graphed main effects plot across clusters and for each cluster separately (Figures A.1.9 and A.1.10). Figure A.1.11 provides the 300 NPI combinations simulated in the four-way scenario analysis. The shaded cells identify NPI combinations with total infections less than 30,600 cases in both Clusters 2 and 6. In addition, we ran an Analysis of Variance (ANOVA) using Minitab 19.2020.1 (Minitab LLC) (Figure A.1.12). The results suggest that social distancing and face mask are the most important NPIs, regardless of cluster type. We found interactions between variables were negligible.

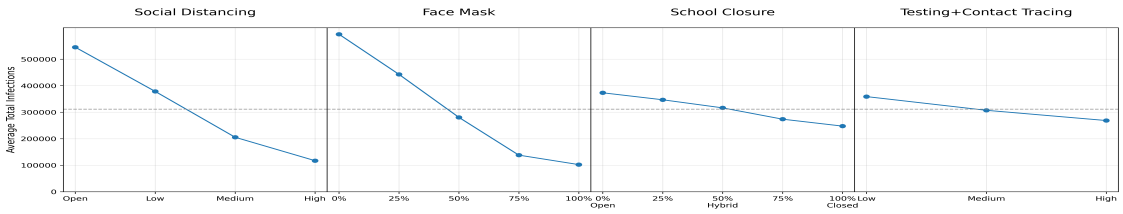


Figure A.1.9: Main effects plot of average total infections in the scenarios of four-way analyses in Clusters 2 and 6, combined. Overall mean marked with lines.

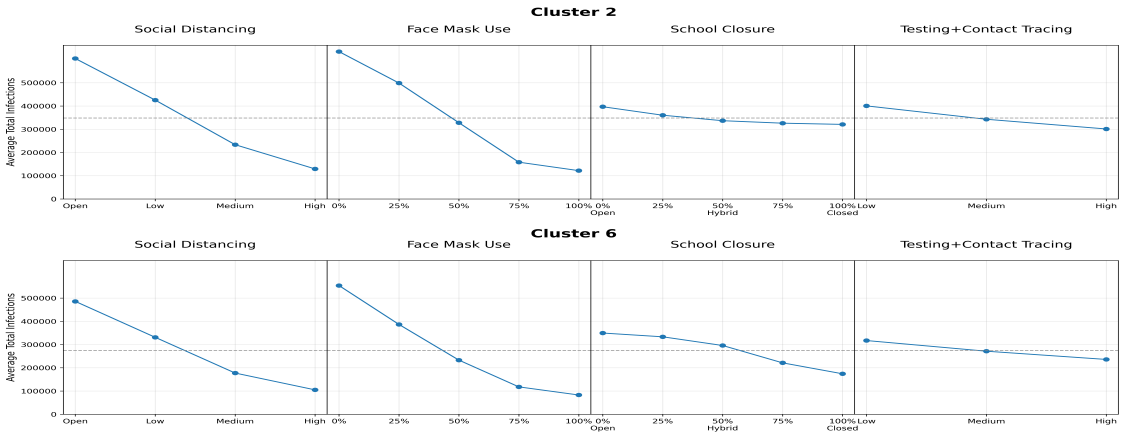


Figure A.1.10: Main effects plot of average total infections in the scenarios of four-way analyses in Clusters 2 and 6, respectively. Overall mean marked with lines.

Face Mask	Social Distancing	School Closure	Testing + Contact Tracing
100%	High	100%	High
100%	High	100%	Medium
100%	High	100%	Low
100%	High	75%	High
100%	High	75%	Medium
100%	High	75%	Low
100%	High	50%	High
100%	High	50%	Medium
100%	High	50%	Low
100%	High	25%	High
100%	High	25%	Medium
100%	High	25%	Low
100%	High	0%	High
100%	High	0%	Medium
100%	High	0%	Low
100%	Medium	100%	High
100%	Medium	100%	Medium
100%	Medium	100%	Low
100%	Medium	75%	High
100%	Medium	75%	Medium
100%	Medium	75%	Low
100%	Medium	50%	High
100%	Medium	50%	Medium
100%	Medium	50%	Low
100%	Medium	25%	High
100%	Medium	25%	Medium
100%	Medium	25%	Low
100%	Medium	0%	High
100%	Medium	0%	Medium
100%	Medium	0%	Low
100%	Low	100%	High
100%	Low	100%	Medium
100%	Low	100%	Low
100%	Low	75%	High
100%	Low	75%	Medium
100%	Low	75%	Low
100%	Low	50%	High
100%	Low	50%	Medium
100%	Low	50%	Low
100%	Low	25%	High
100%	Low	25%	Medium
100%	Low	25%	Low
100%	Low	0%	High
100%	Low	0%	Medium
100%	Low	0%	Low
100%	Open	100%	High
100%	Open	100%	Medium
100%	Open	100%	Low
100%	Open	75%	High
100%	Open	75%	Medium
100%	Open	75%	Low
100%	Open	50%	High
100%	Open	50%	Medium
100%	Open	50%	Low
100%	Open	25%	High
100%	Open	25%	Medium
100%	Open	25%	Low
100%	Open	0%	High
100%	Open	0%	Medium
100%	Open	0%	Low

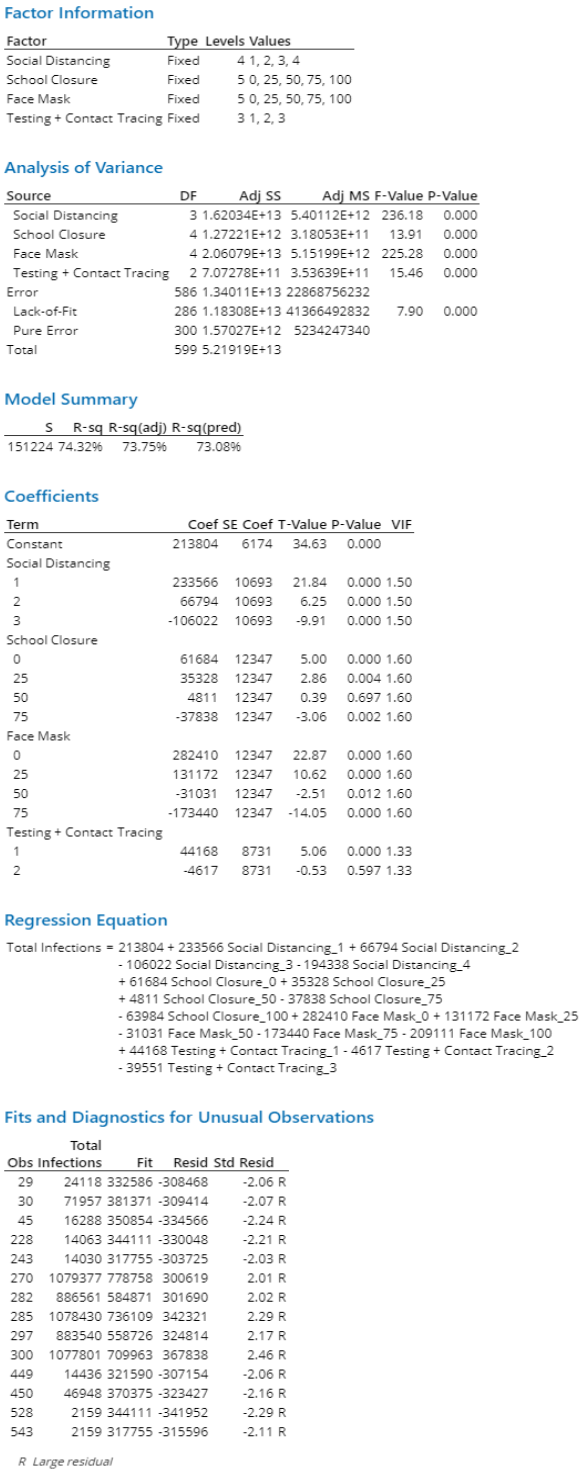
Face Mask	Social Distancing	School Closure	Testing + Contact Tracing
75%	High	100%	High
75%	High	100%	Medium
75%	High	100%	Low
75%	High	75%	High
75%	High	75%	Medium
75%	High	75%	Low
75%	High	50%	High
75%	High	50%	Medium
75%	High	50%	Low
75%	High	25%	High
75%	High	25%	Medium
75%	High	25%	Low
75%	High	0%	High
75%	High	0%	Medium
75%	High	0%	Low
75%	Medium	100%	High
75%	Medium	100%	Medium
75%	Medium	100%	Low
75%	Medium	75%	High
75%	Medium	75%	Medium
75%	Medium	75%	Low
75%	Medium	50%	High
75%	Medium	50%	Medium
75%	Medium	50%	Low
75%	Medium	25%	High
75%	Medium	25%	Medium
75%	Medium	25%	Low
75%	Medium	0%	High
75%	Medium	0%	Medium
75%	Medium	0%	Low
75%	Low	100%	High
75%	Low	100%	Medium
75%	Low	100%	Low
75%	Low	75%	High
75%	Low	75%	Medium
75%	Low	75%	Low
75%	Low	50%	High
75%	Low	50%	Medium
75%	Low	50%	Low
75%	Low	25%	High
75%	Low	25%	Medium
75%	Low	25%	Low
75%	Low	0%	High
75%	Low	0%	Medium
75%	Low	0%	Low
75%	Open	100%	High
75%	Open	100%	Medium
75%	Open	100%	Low
75%	Open	75%	High
75%	Open	75%	Medium
75%	Open	75%	Low
75%	Open	50%	High
75%	Open	50%	Medium
75%	Open	50%	Low
75%	Open	25%	High
75%	Open	25%	Medium
75%	Open	25%	Low
75%	Open	0%	High
75%	Open	0%	Medium
75%	Open	0%	Low

Face Mask	Social Distancing	School Closure	Testing + Contact Tracing
50%	High	100%	High
50%	High	100%	Medium
50%	High	100%	Low
50%	High	75%	High
50%	High	75%	Medium
50%	High	75%	Low
50%	High	50%	High
50%	High	50%	Medium
50%	High	50%	Low
50%	High	25%	High
50%	High	25%	Medium
50%	High	25%	Low
50%	High	0%	High
50%	High	0%	Medium
50%	High	0%	Low
50%	Medium	100%	High
50%	Medium	100%	Medium
50%	Medium	100%	Low
50%	Medium	75%	High
50%	Medium	75%	Medium
50%	Medium	75%	Low
50%	Medium	50%	High
50%	Medium	50%	Medium
50%	Medium	50%	Low
50%	Medium	25%	High
50%	Medium	25%	Medium
50%	Medium	25%	Low
50%	Medium	0%	High
50%	Medium	0%	Medium
50%	Medium	0%	Low
50%	Low	100%	High
50%	Low	100%	Medium
50%	Low	100%	Low
50%	Low	75%	High
50%	Low	75%	Medium
50%	Low	75%	Low
50%	Low	50%	High
50%	Low	50%	Medium
50%	Low	50%	Low
50%	Low	25%	High
50%	Low	25%	Medium
50%	Low	25%	Low
50%	Low	0%	High
50%	Low	0%	Medium
50%	Low	0%	Low
50%	Open	100%	High
50%	Open	100%	Medium
50%	Open	100%	Low
50%	Open	75%	High
50%	Open	75%	Medium
50%	Open	75%	Low
50%	Open	50%	High
50%	Open	50%	Medium
50%	Open	50%	Low
50%	Open	25%	High
50%	Open	25%	Medium
50%	Open	25%	Low
50%	Open	0%	High
50%	Open	0%	Medium
50%	Open	0%	Low

Face Mask	Social Distancing	School Closure	Testing + Contact Tracing
25%	High	100%	High
25%	High	100%	Medium
25%	High	100%	Low
25%	High	75%	High
25%	High	75%	Medium
25%	High	75%	Low
25%	High	50%	High
25%	High	50%	Medium
25%	High	50%	Low
25%	High	25%	High
25%	High	25%	Medium
25%	High	25%	Low
25%	High	0%	High
25%	High	0%	Medium
25%	High	0%	Low
25%	Medium	100%	High
25%	Medium	100%	Medium
25%	Medium	100%	Low
25%	Medium	75%	High
25%	Medium	75%	Medium
25%	Medium	75%	Low
25%	Medium	50%	High
25%	Medium	50%	Medium
25%	Medium	50%	Low
25%	Medium	25%	High
25%	Medium	25%	Medium
25%	Medium	25%	Low
25%	Medium	0%	High
25%	Medium	0%	Medium
25%	Medium	0%	Low
25%	Low	100%	High
25%	Low	100%	Medium
25%	Low	100%	Low
25%	Low	75%	High
25%	Low	75%	Medium
25%	Low	75%	Low
25%	Low	50%	High
25%	Low	50%	Medium
25%	Low	50%	Low
25%	Low	25%	High
25%	Low	25%	Medium
25%	Low	25%	Low
25%	Low	0%	High
25%	Low	0%	Medium
25%	Low	0%	Low
25%	Open	100%	High
25%	Open	100%	Medium
25%	Open	100%	Low
25%	Open	75%	High
25%	Open	75%	Medium
25%	Open	75%	Low
25%	Open	50%	High
25%	Open	50%	Medium
25%	Open	50%	Low
25%	Open	25%	High
25%	Open	25%	Medium
25%	Open	25%	Low
25%	Open	0%	High
25%	Open	0%	Medium
25%	Open	0%	Low

Face Mask	Social Distancing	School Closure	Testing + Contact Tracing
0%	High	100%	High
0%	High	100%	Medium
0%	High	100%	Low
0%	High	75%	High
0%	High	75%	Medium
0%	High	75%	Low
0%	High	50%	High
0%	High	50%	Medium
0%	High	50%	Low
0%	High	25%	High
0%	High	25%	Medium
0%	High	25%	Low
0%	High	0%	High
0%	High	0%	Medium
0%	High	0%	Low
0%	Medium	100%	High
0%	Medium	100%	Medium
0%	Medium	100%	Low
0%	Medium	75%	High
0%	Medium	75%	Medium
0%	Medium	75%	Low
0%	Medium	50%	High
0%	Medium	50%	Medium
0%	Medium	50%	Low
0%	Medium	25%	High
0%	Medium	25%	Medium
0%	Medium	25%	Low
0%	Medium	0%	High
0%	Medium	0%	Medium
0%	Medium	0%	Low
0%	Low	100%	High
0%	Low	100%	Medium
0%	Low	100%	Low
0%	Low	75%	High
0%	Low	75%	Medium
0%	Low	75%	Low
0%	Low	50%	High
0%	Low	50%	Medium
0%	Low	50%	Low
0%	Low	25%	High
0%	Low	25%	Medium
0%	Low	25%	Low
0%	Low	0%	High
0%	Low	0%	Medium
0%	Low	0%	Low
0%	Open	100%	High
0%	Open	100%	Medium
0%	Open	100%	Low
0%	Open	75%	High
0%	Open	75%	Medium
0%	Open	75%	Low
0%	Open	50%	High
0%	Open	50%	Medium
0%	Open	50%	Low
0%	Open	25%	High
0%	Open	25%	Medium
0%	Open	25%	Low
0%	Open	0%	High
0%	Open	0%	Medium
0%	Open	0%	Low

Figure A.1.11: Results of four-way scenario analysis of 300 NPI combinations. Yellow colored cells satisfy Washington State's objective of 30,600 total infections for both Clusters 2 and 6.



Factor Information

Factor	Type	Levels	Values
Social Distancing	Fixed	4	1, 2, 3, 4
School Closure	Fixed	5	0, 25, 50, 75, 100
Face Mask	Fixed	5	0, 25, 50, 75, 100
Testing + Contact Tracing	Fixed	3	1, 2, 3

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Social Distancing	3	1.62034E+13	5.40112E+12	236.18	0.000
School Closure	4	1.27221E+12	3.18053E+11	13.91	0.000
Face Mask	4	2.06079E+13	5.15199E+12	225.28	0.000
Testing + Contact Tracing	2	7.07278E+11	3.53639E+11	15.46	0.000
Error	586	1.34011E+13	22868756232		
Lack-of-Fit	286	1.18308E+13	41366492832	7.90	0.000
Pure Error	300	1.57027E+12	5234247340		
Total	599	5.21919E+13			

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
151224	74.32%	73.75%	73.08%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	213804	6174	34.63	0.000	
Social Distancing					
1	233566	10693	21.84	0.000	1.50
2	66794	10693	6.25	0.000	1.50
3	-106022	10693	-9.91	0.000	1.50
School Closure					
0	61684	12347	5.00	0.000	1.60
25	35328	12347	2.86	0.004	1.60
50	4811	12347	0.39	0.697	1.60
75	-37838	12347	-3.06	0.002	1.60
Face Mask					
0	282410	12347	22.87	0.000	1.60
25	131172	12347	10.62	0.000	1.60
50	-31031	12347	-2.51	0.012	1.60
75	-173440	12347	-14.05	0.000	1.60
Testing + Contact Tracing					
1	44168	8731	5.06	0.000	1.33
2	-4617	8731	-0.53	0.597	1.33

Regression Equation

Total Infections = 213804 + 233566 Social Distancing_1 + 66794 Social Distancing_2
- 106022 Social Distancing_3 - 194338 Social Distancing_4
+ 61684 School Closure_0 + 35328 School Closure_25
+ 4811 School Closure_50 - 37838 School Closure_75
- 63984 School Closure_100 + 282410 Face Mask_0 + 131172 Face Mask_25
- 31031 Face Mask_50 - 173440 Face Mask_75 - 209111 Face Mask_100
+ 44168 Testing + Contact Tracing_1 - 4617 Testing + Contact Tracing_2
- 39551 Testing + Contact Tracing_3

Fits and Diagnostics for Unusual Observations

Obs	Total Infections	Fit	Resid	Std Resid	
29	24118	332586	-308468	-2.06	R
30	71957	381371	-309414	-2.07	R
45	16288	350854	-334566	-2.24	R
228	14063	344111	-330048	-2.21	R
243	14030	317755	-303725	-2.03	R
270	1079377	778758	300619	2.01	R
282	886561	584871	301690	2.02	R
285	1078430	736109	342321	2.29	R
297	883540	558726	324814	2.17	R
300	1077801	709963	367838	2.46	R
449	14436	321590	-307154	-2.06	R
450	46948	370375	-323427	-2.16	R
528	2159	344111	-341952	-2.29	R
543	2159	317755	-315596	-2.11	R

R Large residual

Figure A.1.12: ANOVA result shows that there is a statistically significant association between total infections and every level of social distancing and face mask, some levels of school closure and testing+contact tracing. There are no significant interactions between interventions.

A.1.7 Phase-based Social Distancing Result

Figure A.1.13 compares phase-based social distancing with the base case at low social distancing and with medium social distancing. In Figures A.1.13 (B) and (D), the degree of social distancing for the phase-based policy changes whereas medium social distancing is constant at 50% and the base case maintains low social distancing at 20%. The impact of phase-based social distancing is shown in Figures A.1.13 (A) and (B).

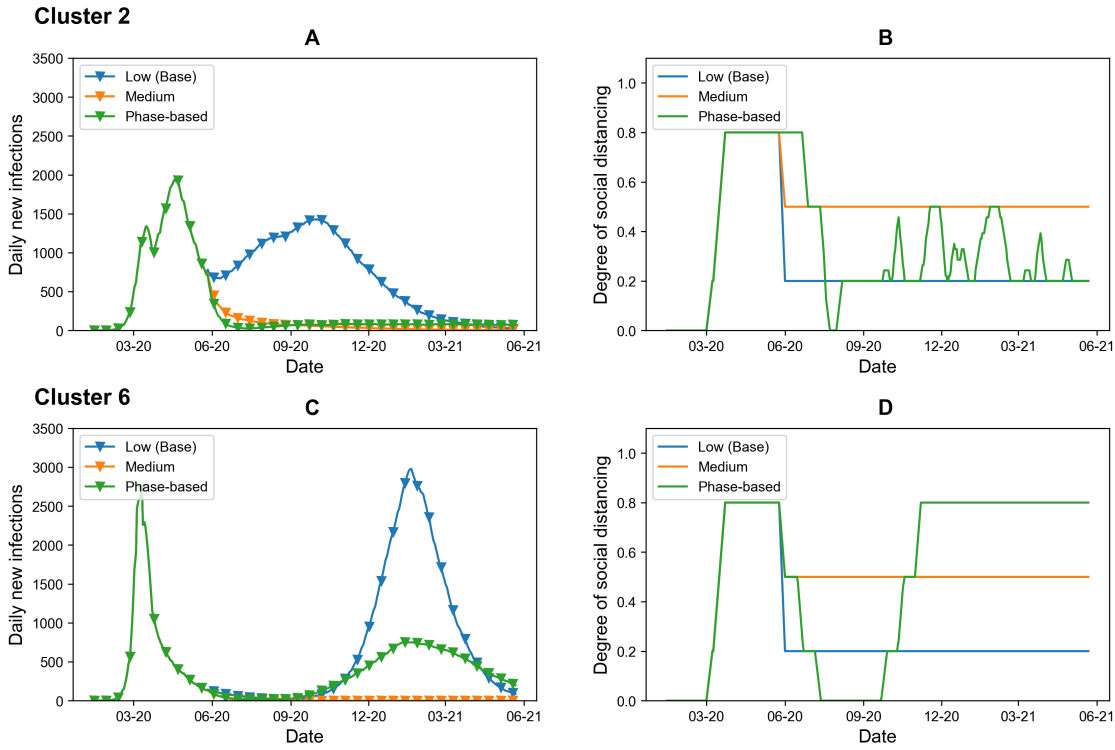


Figure A.1.13: Daily new infections and social distancing level in phase-based social distancing compared to the base case. (A) and (B) indicate changes in daily new infections when low, medium, or phase-based social distancing is enacted since June 1 2020 for Clusters 2 and 6, respectively. (C) and (D) indicate changes in social distancing level for Clusters 2 and 6, respectively.

A.2 Appendix of Section 2.3, Impact of nonpharmaceutical interventions and vaccination with viral variants and waning immunity

A.2.1 COVID-19 Model Specification and Data Input with Viral Variants and Waning Immunity

Population parameters

The synthetic population used in this model is based on U.S. Census data from King County [51]. The total population in the dataset is 1,893,508 individuals, distributed across 789,230 households. The age distribution of the population is described in Figure A.2.14 with a mean value of 36.78 years old. The household size distribution is described in Figure A.2.15 with a mean size of 2.4.

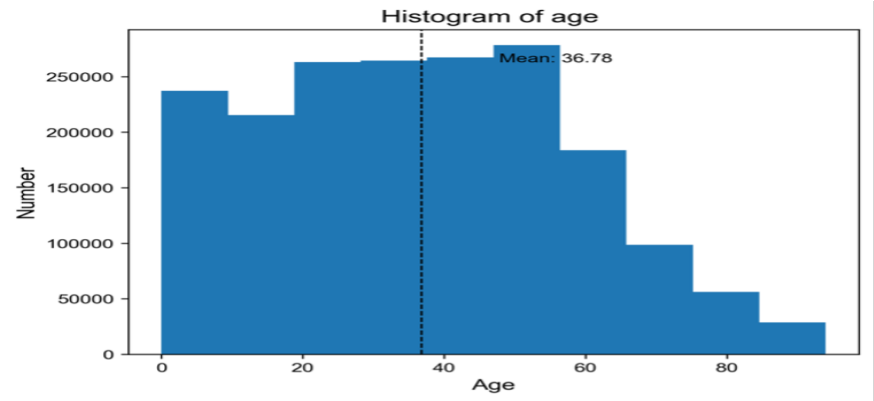


Figure A.2.14: Histogram of age in King County, WA

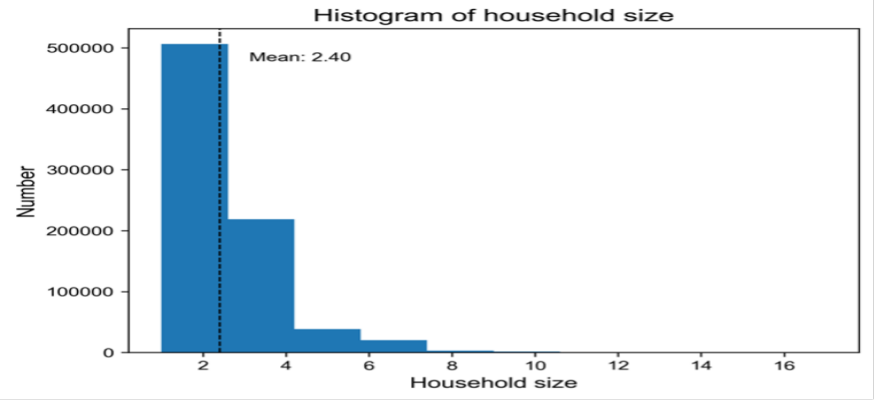
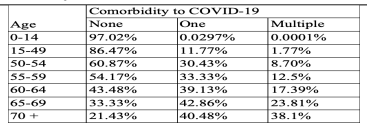
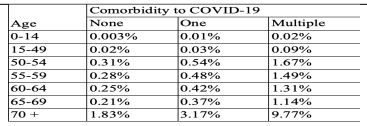


Figure A.2.15: Histogram of household size in King County, WA

Disease parameters

Table A.2.7 provides disease parameters for the COVID-19 model.

Disease Parameters	Value	Reference
Reproduction number (R_0) of original Wuhan strain[1]	2-4	[6]
Infectivity of SARS-CoV-2 original Wuhan strain	9.8%	Calibrated[2]
Immune evasiveness and immune protection against death of original Wuhan strain	See Appendix Section A.2.1	
Incubation period distribution	Lognormal with median of 5.1 days, dispersion of 1.52 days	[28]
Symptomatic period	13 days	[43]
Pre-symptomatic period	2 days	[20]
Infectious period	15 days	[20, 53]
Daily infectivity for 0-14 days	0.5, 0.69, 0.94, 0.8, 0.61, 0.47, 0.35, 0.27, 0.2, 0.16, 0.12, 0.09, 0.07, 0.05,0.04	[3, 20, 53]
Infectiousness of asymptomatic individuals relative to symptomatic	0.75	[5]
Daily household contact probability	85.6%	Calibrated[2]
Daily contact rate in neighborhood	1.188	Calibrated[2]
Daily contact rate in school[3]	3.508	Calibrated[2]
Daily contact rate in workplace[3]	1.745	Calibrated[2]
Probability of asymptomatic infection by age	0-9: 0.71, 10-19: 0.79, 20-29: 0.73, 30-39: 0.67, 40-49: 0.6, 50-59: 0.51, 60-69: 0.37, 70+: 0.31	[11]
Days from infection to death[4]	15 days	[6]
Population distribution that has comorbidity to COVID-19 by age		[10]
Infection fatality ratio by age and comorbidity[5]		[1, 6, 10]
Background annual death rate by age and gender		[16]

Age	Comorbidity to COVID-19		
	None	One	Multiple
0-14	97.02%	0.0297%	0.0001%
15-49	86.47%	11.77%	1.77%
50-54	60.87%	30.43%	8.70%
55-59	54.17%	33.33%	12.5%
60-64	43.48%	39.13%	17.39%
65-69	33.33%	42.86%	23.81%
70 +	21.43%	40.48%	38.1%

Age	Comorbidity to COVID-19		
	None	One	Multiple
0-14	0.003%	0.01%	0.02%
15-49	0.02%	0.03%	0.09%
50-54	0.31%	0.54%	1.67%
55-59	0.28%	0.48%	1.49%
60-64	0.25%	0.42%	1.31%
65-69	0.21%	0.37%	1.14%
70 +	1.83%	3.17%	9.77%

Age	Female	Male	Age	Female	Male
Under 1 year	0.5%	0.61%	45-54	0.3%	0.49%
1-4	0.02%	0.03%	55-64	0.67%	1.12%
5-14	0.01%	0.01%	65-74	1.42%	2.2%
15-24	0.04%	0.10%	75-84	3.79%	5.16%
25-34	0.08%	0.18%	85+	12.87%	14.5%
35-44	0.14%	0.25%			

[1] The R_0 range is used as a target in calibration. See Appendix Section 'Parameter Calibration'

[2] The value was the result of the calibration. See Appendix Section 'Paramber Calibration'

[3] Workplace contacts include office contacts which is a smaller mixing network with twice of the workplace contact rate per day. Same policy is applied for school which includes classroom contact

[4] Calculated from summing 12.9 days from symptom onset to death for 65 or more years plus 2 days of pre-symptomatic period [6]

[5] Calculated from using probability of death by age [6], probability of comorbidity by age [10], and odds ratios for infection fatality by comorbidity [1]

Table A.2.7: COVID-19 Disease Parameters

Variant parameters

Table A.2.8 provides the SARS-CoV-2 variant parameters for the Alpha, Delta, B.1.1.529 Omicron variants and the twelve mutation variants, relative to the previous strain. For example, the Alpha variant is 50% more infectious, 51% more severe, and has the same immune protection compared to the original Wuhan Strain. One of the twelve mutation variants is introduced on June 4, 2022, and then this new variant is assumed to mutate every 6 months. Thus, the first generation of the new variant is imported to the society from 06/04/2022 to 12/04/2022. The second generation of the new variant is imported from 12/05/2022 to 06/03/2023 and has the parameters relative to the first generation of the new variant. The third and last generation of the new mutated generation is imported from 06/04/2023 to 12/31/2023. At every import day, we assume that half of the imports are new SARS-CoV-2 variants and the other half of the imports are the original Wuhan strain. See Figure A.2.16 for the variant introduction timeline.

For ease of explanation, we name the three immune evasion parameter settings as "pessimistic" that includes S1, S4, S7, S10, "neutral" that includes S2, S5, S8, S11, and "optimistic" immune evasion that includes S3, S6, S9, and S12. In the optimistic immune evasion scenarios, we assume that a pan-coronavirus vaccine is always available for all strains, and the vaccine effects always show the same level as against pre-Omicron strains. Appendix Section 'Immune Response Level Equations' provides the detailed immune-related equations.

Variants	Changes in infectivity compared to previous variant (References)	Changes in disease severity compared to previous variant (References)	Changes in immune evasion compared to previous variant	Important dates in mm/dd/yy (References)
Alpha	50% more infectious ([18])	51% more severe ([15])	No difference	01/01/21 to 04/02/21 ([48])
Delta	60% more infectious ([12])	54% more severe ([15])	No difference	04/03/21 to 12/03/21 ([48])
Omicron (B.1.1.529)	40% more infectious ([23])	88% less severe ([42])	Lower	12/04/21 to 06/03/22 ([48])
Mutation Scenarios				
S1	50% more infectious	Same	Pessimistic immune evasion[1]	6/4/2022 to 12/3/2022 (First generation), 12/4/2022 to 6/4/2023 (Second generation), 6/5/2023 to 12/31/2023 (Third generation)
S2			Neutral immune evasion[2]	
S3			Optimistic immune evasion[3]	
S4		50% less severe	Pessimistic immune evasion[1]	
S5			Neutral immune evasion[2]	
S6			Optimistic immune evasion[3]	
S7	Same	Same	Pessimistic immune evasion[1]	
S8			Neutral immune evasion[2]	
S9			Optimistic immune evasion[3]	
S10		50% less severe	Pessimistic immune evasion[1]	
S11			Neutral immune evasion[2]	
S12			Optimistic immune evasion[3]	

[1] 25% more immune evasive that is acquired from vaccination or previous infection
[2] Same immune evasiveness as B.1.1.529 Omicron
[3] Same immune response as B.1.1.529 Omicron when it is acquired from previous infection but enhanced immune response when it is acquired from vaccination

Table A.2.8: SARS-CoV-2 Variant Parameters

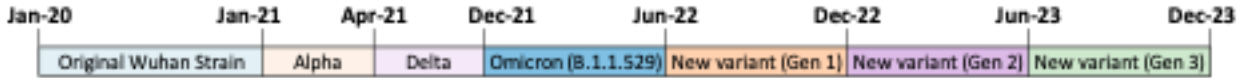


Figure A.2.16: Timeline for variant introduction

Immune response level equations

Immune evasion equation Immunity acquired from vaccination (I_v), previous infection (I_p), or both (I_h) reduces the probability of getting infected. The level of immune evasion is calculated as (1 – immune protection against infection (I_v, I_p, or I_h)). The immune protection against infection is a function of last vaccination date d_v, last infection date d_p, and variant type x. The linear regression equations for the pre-Omicron strains and Omicron B.1.1.529 are given in Table A.2.9. If an individual has both vaccination and previous infection history, the hybrid immunity protection (I_h) is applied which multiplies immune protection from vaccination and previous infection. See Appendix Section 'Transmission Equation' on how it impacts the transmission probability.

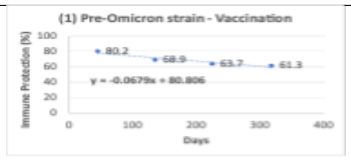
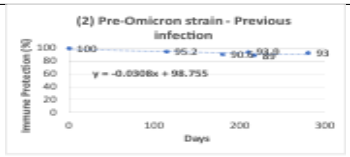
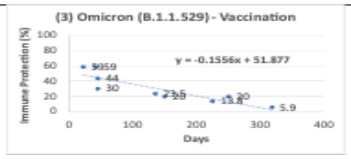
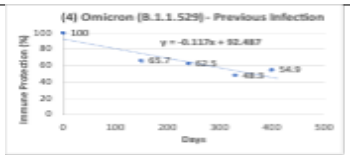
Immune protection against **Infection**		**Immune History**		
		Vaccination (I_v) d_v = days since last vaccination date	Previous Infection (I_p) d_p = days since last infection date	Hybrid (I_h) d_v = days since last vaccination date d_p = days since last infection date
Variant	Pre-Omicron strains (Original Wuhan, Alpha, and Delta)			$I_h^{\text{Pre-O}}(d_v, d_p)$ = $1-(1-I_v^{\text{Pre-O}}(d_v))\cdot(1-I_p^{\text{Pre-O}}(d_p))$
	Equation	$I_v^{\text{Pre-O}}(d_v)$ = max $((-0.07d_v+80.81)/100,0)$	$I_p^{\text{Pre-O}}(d_p)$ = max$((-0.031d_p+98.76)/100,0)$	
	References	[39]	[9, 37]	
	Omicron (B.1.1.529)			$I_h^{\text{B.1.1.529}}(d_v, d_p)$ = 1-$(1-I_v^{\text{B.1.1.529}}(d_v))\cdot(1-I_p^{\text{B.1.1.529}}(d_p))$
	Equation	$I_v^{\text{B.1.1.529}}(d_v)$ = max((-0.156d_v +51.88)/100,0)	$I_p^{\text{B.1.1.529}}(d_p)$)= max((-0.117$d_p$ +92.49)/100,0)	
	References	[19, 39, 41]	[2]	

Table A.2.9: Immune protection equation against infection

When one of the twelve new variants is introduced into the model, its immune evasiveness is relative to the previous variant, i.e., B.1.1.529 Omicron. For example, S1 as in Table A.2.8, is 25% more immune evasive than the previous variant, so the regression equations for the first generation

of S1 (I_v^{n1}, I_p^{n1}, or I_h^{n1}) is 75% of the level of B.1.1.529 Omicron, $I_v^{n1}(d_v) = 0.75 \cdot I_v^{\text{B.1.1.529}}(d_v)$, $I_p^{n1}(d_p) = 0.75 \cdot I_p^{\text{B.1.1.529}}(d_p)$, and $I_h^{n1}(d_v, d_p) = 1 - (1 - I_v^{\text{n1}}(d_v)) \cdot (1 - I_p^{n1}(d_p))$. When the virus mutates after 6 months, the second generation is now n2, and the equations are $I_v^{n2}(d_v) = 0.75 \cdot I_v^{\text{n1}}(d_v)$, $I_p^{n2}(d_p) = 0.75 \cdot I_p^{\text{n1}}(d_p)$, and $I_h^{n2}(d_v, d_p) = 1 - (1 - I_v^{\text{n2}}(d_v)) \cdot (1 - I_p^{n2}(d_p))$. The same logic is applied with the next generation of the mutation n3.

When a pan-coronavirus vaccine is available, as in the "optimistic" mutation scenarios S3, S6, S9, and S12, the immune protection from vaccination is the same for all generations of the new variant, n1, n2, n3, $I_v^{\text{n1,n2,n3}}(d_v) = I_v^{\text{Pre-O}}(d_v)$. The level of immune protection for the other variants that exist are set to Pre-Omicron levels, i.e, $I_v^x(d_v) = I_v^{\text{Pre-O}}(d_v)$ for x= Alpha, Delta, B.1.1.529 Omicron.

While some studies [33, 36] assume high immune protection if previously infected with the same strain type, we do not consider cross-immunity and distinguish whether the previous strain type and the current strain type are the same. Pre-Omicron strains do not need to be distinguished because the variants have strong immunity to each other [36]. However, Omicron strains are highly stealthy immune evaders, so someone infected with Omicron exhibits limited immune effects against Omicrons strains [52]. Since Omicron strains are currently the dominant strains that are highly contagious, we assume that new variants exhibit similar characteristics on the immune evasiveness to the B.1.1.529 Omicron. Thus, we do not assign higher immune protection when individuals come in contact with the same viral type.

Immune protection against death Immunity acquired from vaccination or previous infection reduces the probability of death. The level of immune protection against death (I_D) is a function of last vaccination date d_v, last infection date d_p, and variant type x. The linear regression equations for the pre-Omicron strains and Omicron B.1.1.529 are given in Table A.2.10. See Appendix Section 'Infection Fatality Equation' on how it impacts the infection fatality equation.

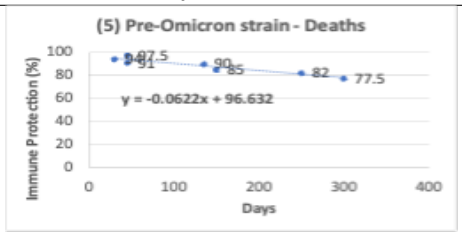
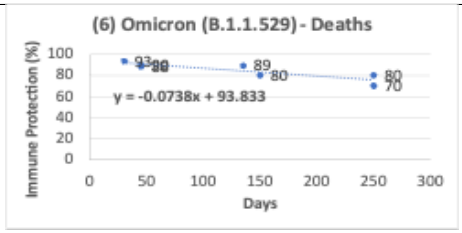
Immune protection against **Death**		**Immune History** (I_D
Variant	Pre-Omicron strains (Original Wuhan, Alpha, and Delta)	
	Equation	$I_D^{\text{Pre-O}}(d_v, d_p)$ = max $((-0.062\min(d_v, d_p) + 96.63)/100, 0)$ where d_v = days since last vaccination date, d_p = days since last infection date
	References	[14, 40]
	Omicron (B.1.1.529)	
	Equation	$I_D^{\text{B.1.1.529}}(d_v, d_p)$ = max $((-0.074\min(d_v, d_p) + 93.83)/100, 0)$ where d_v = days since last vaccination date, d_p = days since last infection date
	References	[14, 41]

Table A.2.10: Immune protection equation against death

When a pan-coronavirus vaccine is available, as in the "optimistic" mutation scenarios S3, S6, S9, and S12, the immune protection from vaccination against death is the same for all generations of the new variant, n1, n2, n3, $I_D^{\text{n1,n2,n3}}(d_v, 0) = I_D^{\text{Pre-O}}(d_v, 0)$. The level of immune protection for the other variants that exist are set to Pre-Omicron levels, i.e, $I_D^x(d_v, 0) = I_D^{\text{Pre-O}}(d_v, 0)$ for $x =$ Alpha, Delta, B.1.1.529 Omicron. In S3, S6, S9, and S12, because the immune protection against death that is acquired from previous infection is at the B.1.1.529 Omicron level, we compute each individual's immune protection from vaccination and previous infection, respectively, and select the maximum value. Thus, $I_D^x(d_v, d_p) = \max(I_D^{\text{Pre-O}}(d_v, 0), I_D^{\text{B.1.1.529}}(0, d_p)$ for $x =$ B.1.1.529, n1, n2, n3.

Timeline, NPI and vaccination parameters

Table A.2.11 provides other parameters for the COVID-19 model, including timeline and NPI parameters. Table A.2.12 provides vaccination parameters.

Timeline Parameters	**Value**	**Reference**
Simulation Period	January 15, 2020 to December 31, 2023	Assumed
Initial infector setting	35 year-old who was infectious from January 15, 2020, and was isolated 5 days later	[21]
Daily imported new infections[1]	January to June 2020: 0 July to September 2020: 0.025 October 2020 to December 2023: 0.558	[35]
Proportion new imports are SARS-CoV-2 variants from January 1, 2021	50%	Assumed
NPI Parameters	**Value**	**Reference**
Total diagnosed cases per 100K population in two weeks to enter phase 2,3,4 on Threshold NPI relaxation	Under 350, 200, 100	[47]
Minimum policy duration on Threshold NPI relaxation	3 weeks	Assumed
Effectiveness of wearing face mask[2]	25%	[13]
Home quarantine duration when contact traced	14 days since last contact with a person who has COVID-19	[7]
Home quarantine percentage when symptomatic without testing	19.5%	Calibrated[3]
Reduction in household contact when home quarantine from March 1, 2020	50% less	Assumed
Delay from onset to testing	2 days	[25]
Delay to receive test results	1 day	[25]
Test sensitivity	0.9	Assumed
Infector's compliance with contact tracing	0.9	Assumed
Contact tracing effectiveness (Household, school, workplace, neighborhood member's compliance with home quarantine and testing when contact traced)	0.9, 0.72, 0.16, 0.10	[27, 29]
Delays in reaching contacts	2 days	[25]

[1] Port of Seattle publishes the number of airline passengers who test positive for COVID-19 every three months from January 2020. Assuming detection rate is 40% [29], we estimated daily infections by multiplying 2.5 (detection rate) and dividing by 90 days. Data on 2021 and 2022 were not available at the time of this study, so we fixed the values at the same level as that of October to December 2020

[2] Assuming people wear surgical masks or cloth masks. Although some studies show higher mask use effectiveness [8], around 50 to 75%, we assumed that the effect of actual mask use is lower because people's mask wearing behavior in a natural environment is not as strict as that of the experimental environment

[3] The value was the result of the calibration. See Appendix Section 'Paramber Calibration'

Table A.2.11: Timeline and NPI parameters

Vaccination Parameters	Value	Reference
Vaccine prioritization policy	WA's phased priority (See below)	[46, 49]
Vaccination willingness by age and vaccination schedule	<table><tr><td></td><td colspan="2">Vaccination schedule</td></tr><tr><td>Age</td><td>Primary (1st and 2nd)</td><td>Booster (3rd)</td></tr><tr><td>0-11</td><td>0.534</td><td>0.128</td></tr><tr><td>12-17</td><td>0.772</td><td>0.374</td></tr><tr><td>18-34</td><td>0.776</td><td>0.413</td></tr><tr><td>35-49</td><td>0.907</td><td>0.588</td></tr><tr><td>50-64</td><td>0.95</td><td>0.694</td></tr><tr><td>65-120</td><td>0.95</td><td>0.902</td></tr></table>	[26][1]
Vaccine dosage policy	Additional dose recipients are prioritized over unvaccinated people on each day	Assumed
Vaccination schedule	See below	
Number of vaccine dosage	7 doses at maximum	Assumed
Vaccine dose interval after previous dose	21 days for 2nd / 6 months for 3rd and more doses	[17]
Vaccine effectiveness	Continuous, preventing infection and severe diseases (See below)	

[1] Values are based on July 6, 2022 values

Table A.2.12: Vaccination parameters

Vaccine prioritization policy We modeled vaccine prioritization policies based on WA's phased policy [49]. The phased priority order is as follows: 1) 65 years or older, or 50 years or older and household size is 4 or more, 2) 50 years or older or essential workers, 3) adults with two or more comorbidities, 4) essential workers younger than 50, 5) essential workers or have one comorbidity, and 6) everyone else.

Vaccine eligibility date by age Individuals are assumed to be eligible for the vaccine if they are ages 16 and older from January 1, 2021, ages 12 and over from May 12, 2021 [45], ages 5 and over from November 1, 2021 [38], and age 6 months and over from June 21, 2022 [24].

Vaccination schedule Our daily dose of vaccines is based on the actual vaccine schedule administered by King County, WA from January to March 2021, after which the vaccine supply is the same as at the end of March 2021 [26]. We adjusted the numbers considering that the synthetic population is based on 2010 census data. Figure A.2.17 shows the simulated vaccination schedule in King County.

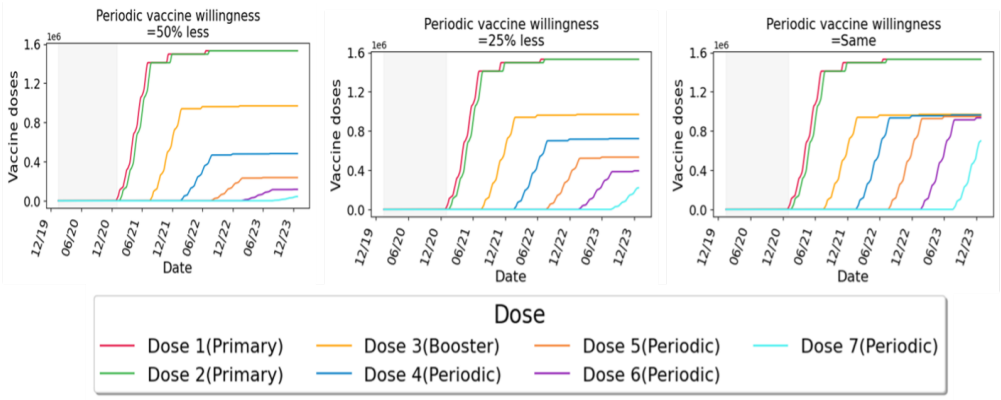


Figure A.2.17: Simulated trends in vaccine administered when periodic vaccine willingness reduces by 50% less, 25% less, or same

Parameter calibration

We calibrated the following unknown model parameters: Infectivity of SARS-CoV-2, daily contact probability between household members, daily contact rates of neighborhood, school, and workplace, and default home quarantine percentage of symptomatic individuals. In the first step, we sampled 1,000 candidate parameter sets using the Latin hypercube method and simulated each parameter set with 200 replications. We initially infected 10 random individuals and simulated for 30 days and identified average secondary infections caused by the initial infectors to calculate R0. Through this process, 55 parameter sets yielded R0 values between 2 and 4, a known range for COVID-19 [6]. Then, we used Minitab 19.2020.1 (Minitab LLC) to identify clusters of the 55 parameter sets using complete-linkage method. This resulted in 8 clusters.

In the second step, we simulated public health interventions in King County from January 15, 2020, through December 31, 2020, by observing the region's sequence of interventions (See Figure A.2.18). We fit compliance to social distancing using IHME's data on WA's social distancing [22]. We assumed that compliance to social distancing is low if changes in mobility is -20 to -30%, medium if the change is -30 to -40%, and high if the change is over -40%. From late August to early October 2020, there was a change in mobility between -16% and -19% in some days, but considering the small differences, the social distance was kept low. Compliance to face mask use was also based on IHME data [22]. We identified the first days that face mask usage exceeds 25%, 50%, and 75%, respectively. School closure was based on King County timeline [44]. Testing and contact tracing were based on Lee et al. [29]. Finally, we selected the cluster with the smallest

difference between the time-weighted sum of 7-day moving average reported deaths during the calibration period.

Figure A.2.19 shows calibration results on reported deaths, hospitalizations, and diagnosed cases. Although reported deaths in the second half of 2020 appear slightly later and higher than the model outcomes, the later peak is likely due to the fact that the time from infection to death continues to increase with advances in medical care, and higher mortality could occur due to the congestion in hospitals that cause excess deaths. As hospitalizations and cases show similar trends to the reported data, we conclude our calibration was appropriate.

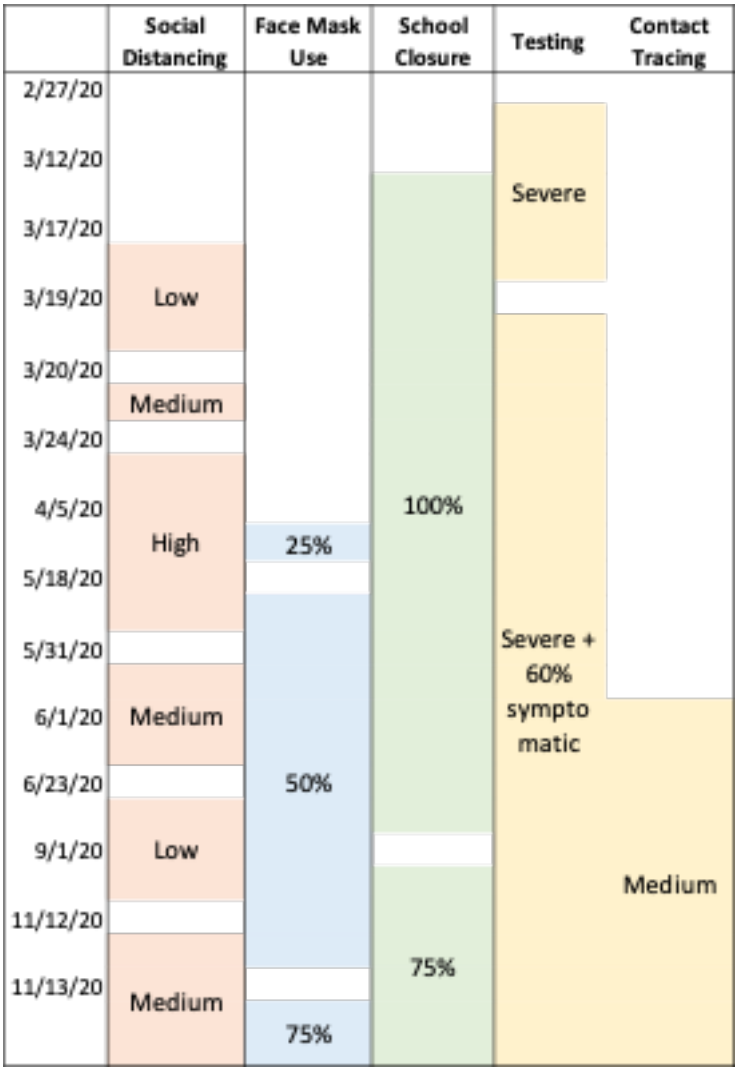


Figure A.2.18: Diagram of simulated interventions from January 15, 2020, to December 31, 2020

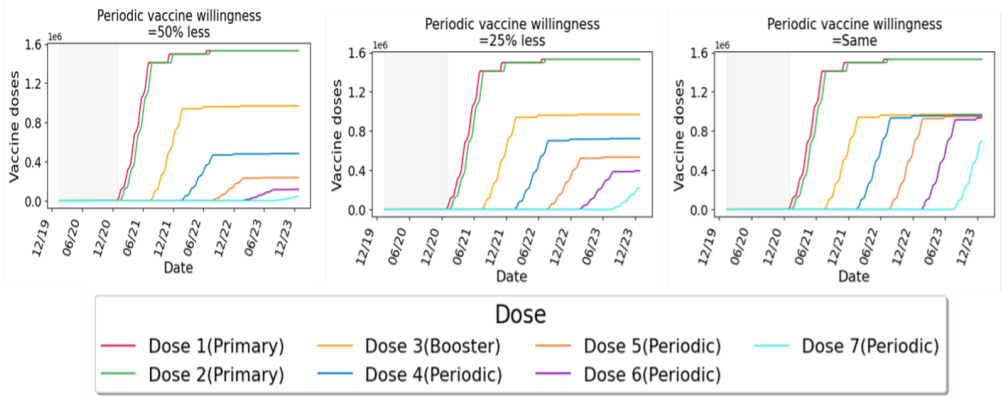


Figure A.2.19: Calibration to reported deaths, hospitalizations, and cases in King County, WA. The 50% confidence interval was obtained by calculating 25% and 75% percentiles at each time step. Hospitalizations reported in King County were based on the date of admission and model results were based on the date of infection. For comparison, we moved the reported hospitalizations to 12 days earlier, and graphed both outcomes based on the date of infection

A.2.2 SARS-CoV-2 Disease Model Equations with with Viral Variants and Waning Immunity

Transmission equation

Household transmission For an infectious person i who is infected from variant type x, we consider his or her household transmission for each dth infectious day. For each susceptible household member j, j becomes infected with Household infection probability (i,j,x,d):

$$\begin{aligned}\text{Household infection probability } (i,j,x,d) = (&\text{daily household contact probability}) \times (1 \text{ - reduction} \\ &\text{in contact if } i \text{ is home quarantined}) \\ &\times (\text{infectivity of x}) \\ &\times \text{daily infectivity}(i,\, d) \times (\text{reduced infectiousness} \\ &\text{if } i \text{ is asymptomatic}) \\ &\times (1\text{-immune protection against infection}(j,x)).\end{aligned}$$

Non-household transmission (neighborhood, a workplace, or a school) If an infectious person i is present at the place p on dth infectious day and is infected from variant type x, the expected

number of contacts is the Contact count (i, x, p, d):

$$\text{Contact count } (i, x, p, d) = (\text{daily contacts per day } (p)) \times (\text{infectivity of } x) \times \text{daily infectivity } (i, d) \times (\text{reduced infectiousness if } i \text{ is asymptomatic}).$$

If Contact count (i,x, p,d) is not integer, stochastic rounding is needed. If Contact count (i,x, p,d) $= k$, then

$$\text{Actual contact count}(i,x,p,d) = Round(k) = \begin{cases} \lfloor k \rfloor \text{ with probability } 1-(k-\lfloor k \rfloor) \\ \lfloor k \rfloor + 1 \text{ with probability } (k-\lfloor k \rfloor) \end{cases}$$

For example, if $k = 2.7$, Actual contact count(i,x, p,d) would be 2 with probability 0.3 and 3 with probability 0.7. We randomly select Actual contact count (i,x, p,d) individuals in place p on day d. For each contactee j in the place p, if j is susceptible and present in the place, j becomes infected with Non-household infection probability (i,j):

$$\text{Non-household infection probability } (i,j) = (1 - \text{effectiveness of wearing face mask if } i \text{ uses mask}) \times (1 - \text{effectiveness of wearing face mask if } j \text{ uses mask}).$$

Symptomatic disease equation

An infected individual i who is infected from variant type x will experience symptoms depending on their age and variant type with Probability of symptomatic disease(i,x):

$$\text{Probability of symptomatic disease}(i,x) = (1\text{-probability of asymptomatic infection by age } (i)) \times (\text{relative disease severity of variant } x)$$

Infection fatality equation

An infected individual i who is infected from variant type x may die from COVID-19 depending on their age, comorbidity, immunity protection and variant type x. The infection fatality equation

is the Probability of death when infected (i,x):

Probability of death when infected (i,x) = (infection fatality ratio of i by age and comorbidity)
$\times$ (relative disease severity of variant x)
$\times$ (1-immune protection against death when infected (i,x))

where

$$\text{immune protection against death when infected } (i,x) = \min\left(\frac{1-I_D^x(i)}{1-I_I^x(i)}, 1\right),$$

$I_D^x(i) = I_D^x(d_v, d_p)$,
$I_I^x(i) = (I_v^x(d_v), I_p^x(d_p)$ or $I_h^x(d_v, d_p))$,
$d_v =$ Individual i's days since last vaccination date,
$d_p =$ Individual i's days since last infections date.

The immune protection against death when infected is based on [30].

A.2.3 Results – Impact of viral mutation on SARS-CoV-2 infections

Figure A.2.20 shows the impact of viral mutation on SARS-CoV-2 infections from January 15, 2020 to December 31, 2023 when periodic vaccination rate is assumed to decrease by 25% for each additional booster vaccine and NPI policy is assumed to be Timeline 1.

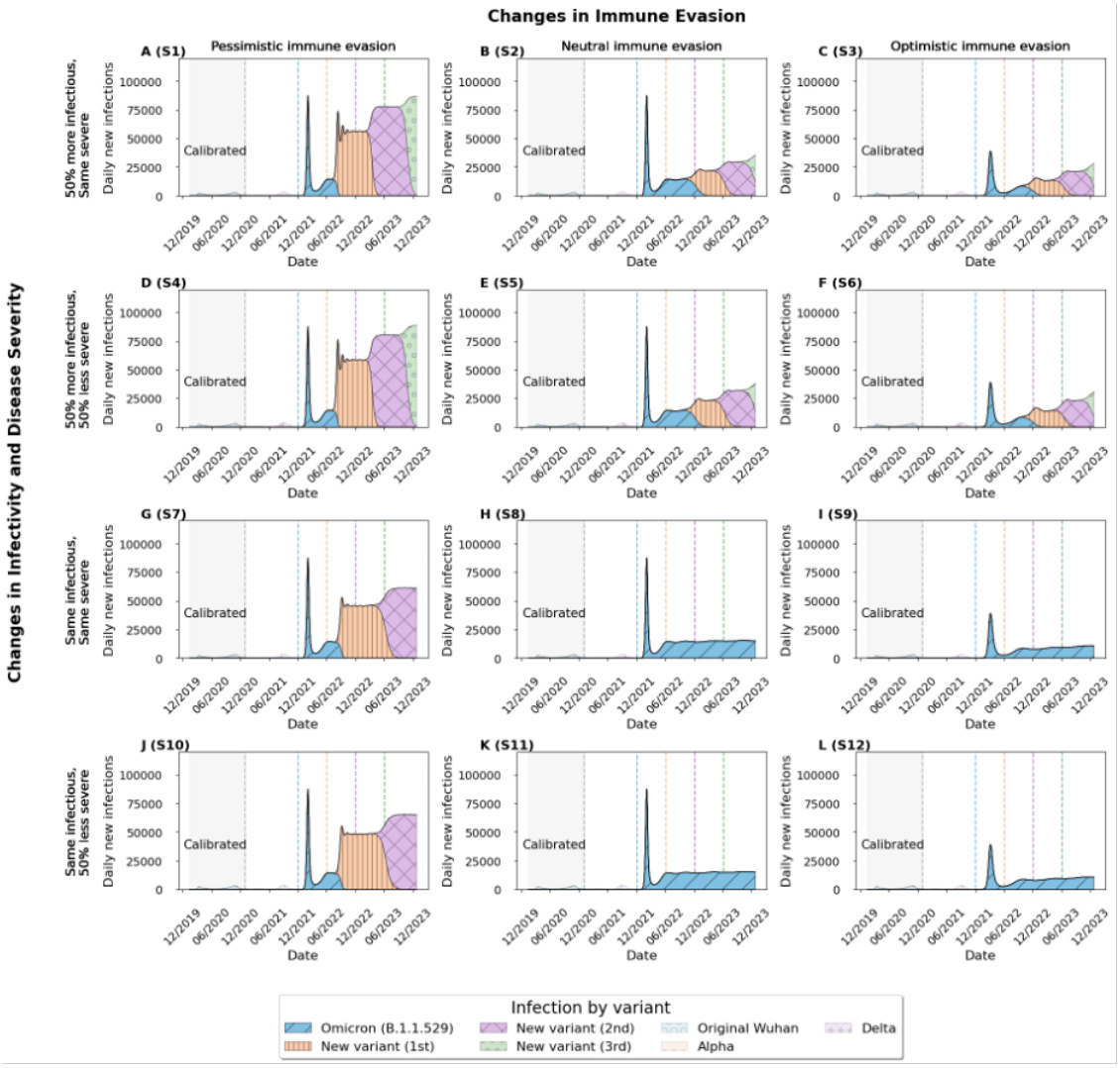


Figure A.2.20: Impact of viral mutation on SARS-CoV-2 infections while varying infectivity and immune evasion from vaccination or previous infection. Virus mutation scenarios are listed in the upper left corner of each graph. Colored lines in each graph indicate the first imported date of each variant. Periodic vaccination rate is assumed to decrease by 25% for each additional booster vaccine and NPI policy is assumed to be Timeline 1

A.2.4 Results – Impact of viral mutation, vaccination, and NPI policy on SARS-CoV-2 daily deaths

Impact of viral mutation and periodic vaccination rate on SARS-CoV-2 daily deaths

Figure A.2.21 illustrates the impact of viral mutation and periodic vaccination rate on SARS-CoV-2 daily deaths from January 15, 2020 to December 31, 2023. NPI policy is assumed to be Timeline 1.

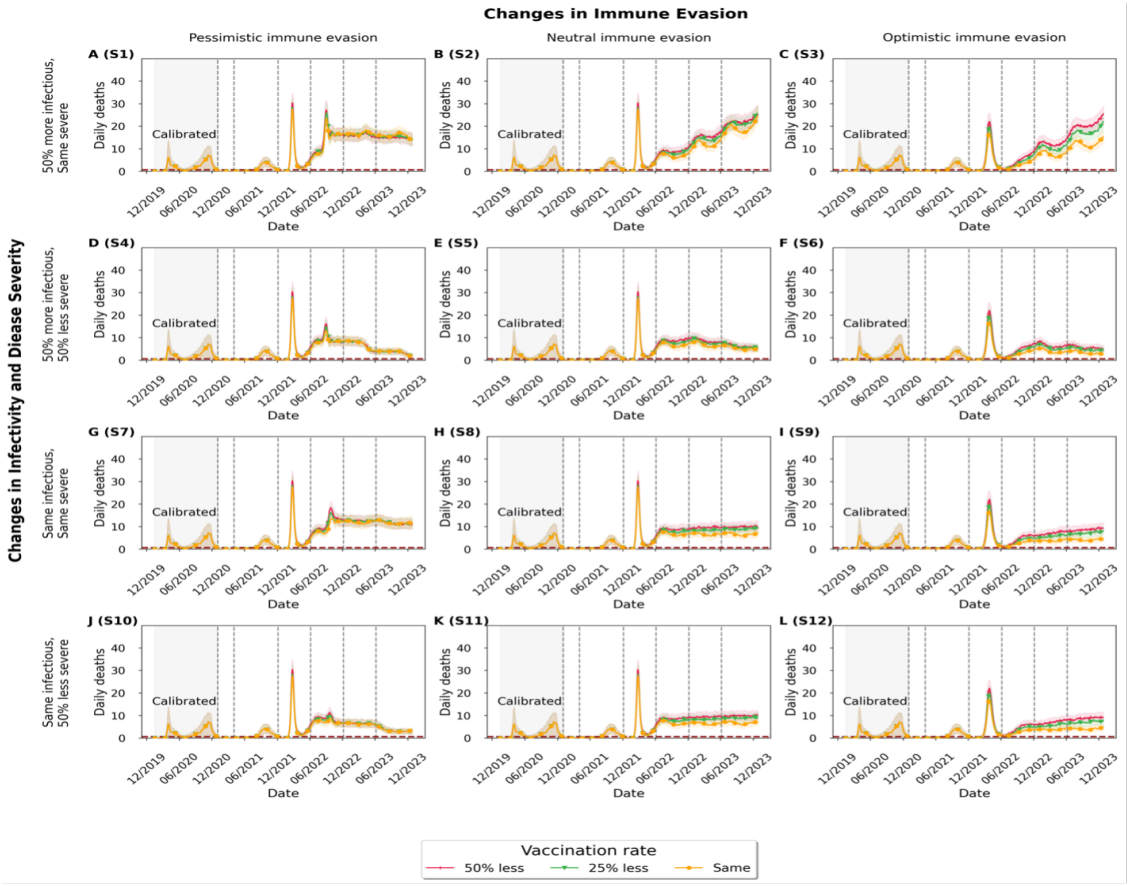


Figure A.2.21: The impact of viral mutation and periodic vaccination willingness policy from January 15, 2020 to December 31, 2023. NPI policy is assumed to follow Timeline 1. Shaded area in each line indicates 25th and 75th percentile values of daily deaths. In each graph, the vertical grey line indicates the first imported date of each variant (the Alpha, the Delta, the B.1.1.529 Omicron, and a new variant that mutates twice more). Red dotted lines represent the number of daily deaths from influenza and pneumonia in Washington state in 2017, which is calculated to be 0.65 deaths per day

Impact of viral mutation and NPI policy on SARS-CoV-2 daily deaths

Figure A.2.22 illustrates the impact of viral mutation and NPI policy on SARS-CoV-2 daily deaths from January 15, 2020 to December 31, 2023. Periodic vaccination rate is assumed to reduce by 25% for each additional dose.

Figure A.2.22: The impact of viral mutation and NPI policy from January 15, 2020 to December 31, 2023. Periodic vaccination rate is assumed to reduce by 25% for each additional dose. Shaded area in each line indicates the first imported date of each variant (the Alpha, the Delta, the B.1.1.529 Omicron, and a new variant that mutates twice more). Red dotted lines represent the number of daily deaths from influenza and pneumonia in Washington state in 2017, which is calculated to be 0.65 deaths per day

A.2.5 Results – Changes in NPI stages when Threshold NPI policy is applied

Figure A.2.23 illustrates the NPI stages in 2021-2023 for the Threshold policy that corresponds to the settings in Figure 2.7 in the manuscript. For example, the Threshold policy in A.2.10A corresponds to the same mutation setting in Figure 2.7A in the manuscript. The NPI stages shown in Figure A.2.23 are averaged by 6-month intervals, where the red shading indicates NPI stage 1, or lockdown, and the blue shading indicates NPI stage 4, or total opening. Before the introduction of new variants (before June 2022), the society fluctuates between NPI stage 2 to 3.

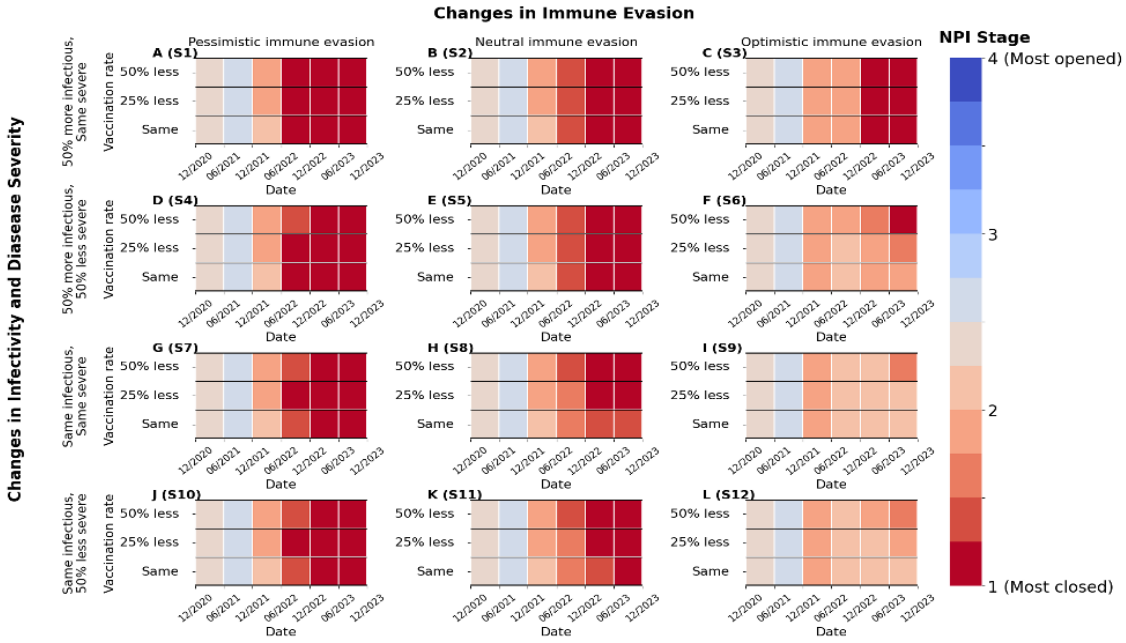


Figure A.2.23: Changes in NPI stages (averaged by 6-month intervals) when Threshold NPI policy is applied with varying mutation scenarios on infectivity, disease severity, and immune protection from vaccination or previous infection

REFERENCES FOR APPENDIX A

1. A FAIR Health White Paper MM. Risk Factors for COVID-19 Mortality among Privately Insured Patients: A Claims Data Analysis. Generic. [Online; accessed 10-August-2023]. 2020. Available from: https://s3.amazonaws.com/media2.fairhealth.org/whitepaper/asset/Risk%20Factors%20for%20COVID-19%20Mortality%20among%20Privately%20Insured%20Patients%20-%20A%20Claims%20Data%20Analysis%20-%20A%20FAIR%20Health%20White%20Paper.pdf

2. Altarawneh HN, Chemaitelly H, Hasan MR, Ayoub HH, Qassim S, AlMukdad S, et al. Protection against the Omicron Variant from Previous SARS-CoV-2 Infection. N Engl J Med 2022; 386:1288–90. DOI: 10.1056/NEJMc2200133

3. Anderson RM, Heesterbeek H, Klinkenberg D, and Hollingsworth TD. How will country-based mitigation measures influence the course of the COVID-19 epidemic? The lancet 2020; 395:931–4

4. Basu A. Estimating The Infection Fatality Rate Among Symptomatic COVID-19 Cases In The United States: Study estimates the COVID-19 infection fatality rate at the US county level. Health Affairs 2020 :10–1377

5. Centers for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. [Online; accessed 15-April-2023]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html

6. Centers for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. [Online; accessed 19-March-2021]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios-archive/planning-ccenarios-2021-03-19.pdf

7. Centers for Disease Control and Prevention. COVID-19: When to Quarantine. [Online; accessed 10-August-2020]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/if-you-are-sick/quarantine.html

8. Centers for Disease Control and Prevention. Science Brief: Community Use of Cloth Masks to Control the Spread of SARS-CoV-2. Generic. [Online; accessed 15-April-2023]. 2021. Available from: https://stacks.cdc.gov/view/cdc/105951

9. Centers for Disease Control and Prevention. Science Brief: SARS-CoV-2 Infection-induced and Vaccine-induced Immunity. Web Page. [Online; accessed 19-October-2021]. 2021. Available from: https://www.cdc.gov/coronavirus/2019-ncov/science/science-briefs/vaccine-induced-immunity.html

10. Clark A, Jit M, Warren-Gash C, Guthrie B, Wang HH, Mercer SW, et al. How many are at increased risk of severe COVID-19 disease? Rapid global, regional and national estimates for 2020. MedRxiv 2020. DOI: 10.1101/2020.04.18.20064774

11. Davies NG, Klepac P, Liu Y, Prem K, Jit M, Pearson CA, et al. Age-dependent effects in the transmission and control of COVID-19 epidemics. Nature Medicine 2020; 26:1205–11. DOI: 10.1038/s41591-020-0962-9

12. Earnest R, Uddin R, Matluk N, Renzette N, Turbett SE, Siddle KJ, et al. Comparative transmissibility of SARS-CoV-2 variants Delta and Alpha in New England, USA. Cell Reports Medicine 2022; 3:100583. DOI: https://doi.org/10.1016/j.xcrm.2022.100583

13. Eikenberry SE, Mancuso M, Iboi E, Phan T, Eikenberry K, Kuang Y, et al. To mask or not to mask: Modeling the potential for face mask use by the general public to curtail the COVID-19 pandemic. Infectious Disease Modelling 2020; 5:293–308. DOI: https://doi.org/10.1016/j.idm.2020.04.001

14. Ferdinands JM, Rao S, Dixon BE, Mitchell PK, DeSilva MB, Irving SA, et al. Waning 2-dose and 3-dose effectiveness of mRNA vaccines against COVID-19–associated emergency department and urgent care encounters and hospitalizations among adults during periods of Delta and Omicron variant predominance—VISION Network, 10 states, August 2021–January 2022. Morbidity and Mortality Weekly Report 2022; 71:255. Available from: https://www.cdc.gov/mmwr/volumes/71/wr/pdfs/mm7107e2-H.pdf

15. Fisman DN and Tuite AR. Evaluation of the relative virulence of novel SARS-CoV-2 variants: a retrospective cohort study in Ontario, Canada. CMAJ 2021; 193:E1619–E1625. DOI: 10.1503/cmaj.211248

16. Flagg LA and Anderson RN. National Vital Statistics Reports. Generic. [Online; accessed 19-March-2021]. 2021. Available from: https://www.cdc.gov/nchs/data/nvsr/nvsr69/nvsr69-14-508.pdf

17. Food Drug Administration. Vaccines and related biological products advisory committee meeting december 10, 2020: FDA briefing document, Pfizer-BioNTech COVID-19 vaccine. Web Page. [Online; Accessed 19-July-2022]. 2020. Available from: https://www.fda.gov/media/144245/download

18. Fort H. A very simple model to account for the rapid rise of the alpha variant of SARS-CoV-2 in several countries and the world. Virus Research 2021; 304:198531

19. Fowlkes AL, Yoon SK, Lutrick K, Gwynn L, Burns J, Grant L, et al. Effectiveness of 2-dose BNT162b2 (Pfizer BioNTech) mRNA vaccine in preventing SARS-CoV-2 Infection among children aged 5–11 years and adolescents aged 12–15 years—PROTECT Cohort, July 2021–February 2022. Morbidity and Mortality Weekly Report 2022; 71:422. Available from: http://dx.doi.org/10.15585/mmwr.mm7111e1external

20. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, et al. Temporal dynamics in viral shedding and transmissibility of COVID-19. Nature medicine 2020; 26:672–5

21. Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, et al. First case of 2019 novel coronavirus in the United States. New England Journal of Medicine 2020

22. Institute for Health Metrics and Evaluation. COVID-19 Projections. [Online; accessed 10-January-2023]. 2020. Available from: https://covid19.healthdata.org/projections

23. Jalali N, Brustad HK, Frigessi A, MacDonald EA, Meijerink H, Feruglio SL, et al. Increased household transmission and immune escape of the SARS-CoV-2 Omicron variant compared to the Delta variant: evidence from Norwegian contact tracing and vaccination data. MedRxiv 2022

24. Kimball S. Covid vaccinations for children under age 5 to begin as early as June 21, White House says. Web Page. [Online; Accessed 14-September-2022]. 2022. Available from: https://www.cnbc.com/2022/06/02/white-house-expects-covid-vaccinations-for-children-under-5-to-begin-as-early-as-june-21.html

25. King County - Department of Health. Case investigation and contact tracing dashboard. Web Page. [Online; accessed 18-July-2021]. 2021. Available from: https://kingcounty.gov/depts/health/covid-19/data/contact-tracing.aspx

26. King County - Department of Health. Summary of COVID vaccination among King County residents. Web Page. [Online; accessed 1-June-2021]. 2021. Available from: https://kingcounty.gov/depts/health/covid-19/data/vaccination.aspx

27. Kucharski A, Klepac P, Conlan A, Kissler S, Tang M, Fry H, et al. Effectiveness of isolation, testing, contact tracing, and physical distancing on reducing transmission of SARS-CoV-2 in different settings: a mathematical modelling study. The Lancet Infectious Diseases 2020 Jun; 20. DOI: 10.1016/S1473-3099(20)30457-6

28. Lauer SA, Grantz KH, Bi Q, Jones FK, Zheng Q, Meredith HR, et al. The incubation period of coronavirus disease 2019 (COVID-19) from publicly reported confirmed cases: estimation and application. Annals of internal medicine 2020; 172:577–82

29. Lee S, Zabinsky ZB, Wasserheit JN, Kofsky SM, and Liu S. COVID-19 Pandemic Response Simulation in a Large City: Impact of Nonpharmaceutical Interventions on Reopening Society. Med Decis Making 2021; 41:419–29. DOI: 10.1177/0272989X211003081

30. Lego V di, Sanchez-Romero M, and Prskawetz A. The impact of COVID-19 vaccines on the Case Fatality Rate: The importance of monitoring breakthrough infections. Int J Infect Dis 2022; 119:178–83. DOI: 10.1016/j.ijid.2022.03.059

31. Liu F, Enanoria WT, Zipprich J, Blumberg S, Harriman K, Ackley SF, et al. The role of vaccination coverage, individual behaviors, and the public health response in the control of measles epidemics: an agent-based simulation for California. BMC public health 2015; 15:447

32. Minitab 18. Interpret the key results for Cluster Variables. [Online; accessed 15-February-2021]. 2019. Available from: https://support.minitab.com/en-us/minitab/18/help-and-how-to/modeling-statistics/multivariate/how-to/cluster-variables/interpret-the-results/key-results/#step-1-examine-the-similarity-and-distance-levels

33. Moghadas SM, Vilches TN, Zhang K, Wells CR, Shoukat A, Singer BH, et al. The Impact of Vaccination on Coronavirus Disease 2019 (COVID-19) Outbreaks in the United States. Clin Infect Dis 2021; 73:2257–64. DOI: 10.1093/cid/ciab079

34. O'Dowd A. Covid-19: Johnson is on back foot over next steps to control pandemic. BMJ 2020; 369. DOI: 10.1136/bmj.m2152

35. Port of Seattle. Confirmed COVID-19 Positive Tests at SEA. Web Page. [Online; accessed 15-March-2021]. 2021. Available from: https://www.portseattle.org/node/13520

36. Reiner B. COVID-19 model update: Omicron and waning immunity. Web Page. 2021. Available from: https://www.healthdata.org/special-analysis/omicron-and-waning-immunity

37. Sheehan MM, Reddy AJ, and Rothberg MB. Reinfection rates among patients who previously tested positive for coronavirus disease 2019: a retrospective cohort study. Clinical Infectious Diseases 2021; 73:1882–6

38. Stankiewicz K. Pfizer director Dr. Scott Gottlieb says Covid vaccine for kids 5 to 11 could come by winter. Web Page. [Online; Accessed 5-September-2021]. 2021. Available from: https://www.cnbc.com/2021/08/30/pfizer-director-dr-scott-gottlieb-on-covid-vaccine-for-kids-5-to-11.html

39. Tseng HF, Ackerson BK, Luo Y, Sy LS, Talarico CA, Tian Y, et al. Effectiveness of mRNA-1273 against SARS-CoV-2 Omicron and Delta variants. Nat Med 2022; 28:1063–71

40. UK Health Security Agency. COVID-19 vaccine surveillance report - Week 11. Web Page. [Online; accessed 14-June-2022]. 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1061532/Vaccine_surveillance_report_-_week_11.pdf

41. UK Health Security Agency. COVID-19 vaccine surveillance report - Week 24. Web Page. [Online; accessed 14-September-2022]. 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1083443/Vaccine-surveillance-report-week-24.pdf

42. Ulloa AC, Buchan SA, Daneman N, and Brown KA. Estimates of SARS-CoV-2 Omicron Variant Severity in Ontario, Canada. JAMA 2022; 327:1286–8. DOI: 10.1001/jama.2022.2274

43. Verity R, Okell LC, Dorigatti I, Winskill P, Whittaker C, Imai N, et al. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The Lancet Infectious Diseases 2020. DOI: 10.1016/S1473-3099(20)30243-7

44. Washington Office of Superintendent of Public Instruction. School Facility ReOpening Survey. Web Page. [Online; Accessed 19-July-2021]. 2021. Available from: https://data.wa.gov/dataset/School-Facility-ReOpening-Survey/9i5d-c2m8

45. Washington State, Department of Health. Everyone 12 and older now eligible for Pfizer-BioNTech COVID-19 vaccine. Web Page. [Online; accessed 5-Sep-2021]. 2021. Available from: https://www.doh.wa.gov/Newsroom/Articles/ID/2780/Everyone-12-and-older-now-eligible-for-Pfizer-BioNTech-COVID-19-vaccine

46. Washington State, Department of Health. Interim COVID-19 Vaccination Plan-October 2020, version 1. Web Page. [Online; accessed 14-Sep-2022]. 2020. Available from: https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/WA-COVID-19-Vaccination-Plan.pdf

47. Washington State, Department of Health. Roadmap to Recovery Metrics. Web Page. [Online; accessed 20-June-2021]. 2021. Available from: https://coronavirus.wa.gov/what-you-need-know/roadmap-recovery-metrics

48. Washington State, Department of Health. SARS-CoV-2 Sequencing and Variants in Washington State. Web Page. [Online; accessed 6-September-2022]. 2022. Available from: https://doh.wa.gov/sites/default/files/2022-02/420-316-SequencingAndVariantsReport.pdf?uid=63179ee18bd2e

49. Washington State, Department of Health. Washington's COVID-19 Vaccine Phases. Web Page. [Online; accessed 1-June-2021]. 2021. Available from: https://www.doh.wa.gov/Portals/1/Documents/1600/coronavirus/VaccinationPhasesInfographic.pdf

50. Washington State, The Office of The Governor. Safe Start Washington – Phased Reopening County-by-County. [Online; accessed 22-October-2020]. 2020. Available from: https://www.governor.wa.gov/sites/default/files/SafeStartPhasedReopening.pdf

51. Wheaton W and RTI International. US Synthetic Population 2010 Version 1.0: Quick Start Guide. [Online; accessed 1-December-2020]. 2014. Available from: https://www.rti.org/impact/rti-us-synthetic-household-population%E2%84%A2

52. Wise J. Covid-19: Omicron infection is poor booster to immunity, study finds. BMJ 2022; 377:o1474. DOI: 10.1136/bmj.o1474

53. Young BE, Ong SWX, Kalimuddin S, Low JG, Tan SY, Loh J, et al. Epidemiologic features and clinical course of patients infected with SARS-CoV-2 in Singapore. JAMA 2020; 323:1488–94

Appendix B

APPENDIX OF CHAPTER 3

Contents

B.1 Input Parameters

This section outlines the input parameters for the coupled dynamics model and vaccination campaign simulations. Table B.1.1 provides the values and references for input model parameters. The descriptions of the parameters are given in part of Table 2 in the main paper.

B.1.1 Physical contact matrix C^{Base}

We obtained the 25×25 physical contact matrix C^{Base}, where C_{ij}^{Base} is the probability that a physical contact occurs from individuals in group i to group j, from two data sources. First, we analyzed a week of mobile phone data (April 1, 2019 - April 8, 2019) provided by SafeGraph [8, 18] to obtain the 5×5 regional contact matrix G^{Base}. Second, we used a literature source [16] to obtain a 5×5 age contact matrix, A^{Base}. We normalized each matrix so their sums equal one. That is, $\sum_{r_x} \sum_{r_y} G_{r_x r_y}^{\text{Base}} = 1$ where r_x and r_y are geographic regions and $\sum_{a_x} \sum_{a_y} A_{a_x a_y}^{\text{Base}} = 1$ where a_x and a_y are age ranges. Assuming that geographic region and age ranges have independent relations, the final contact matrix C^{Base} was calculated using the Kronecker product of G^{Base} and A^{Base}. Figure B.1.1 provides the contact matrices.

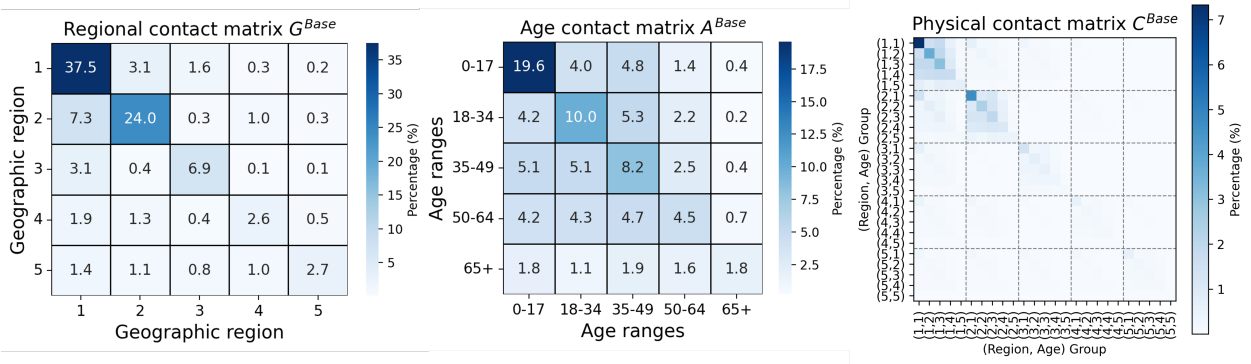


Figure B.1.1: Physical contact matrix (C^{Base}) in King County, obtained by taking a Kronecker product of the regional contact matrix G^{Base} and age contact matrix A^{Base}.

B.1.2 Opinion sharing matrix O^{Base}

Data on 25×25 opinion sharing matrix O^{Base}, where O_{ij}^{Base} is defined as the probability that an opinion sharing occurs from individuals in group i to group j, was unavailable by age and geography level. Thus, in the base case, we assumed that O^{Base} generally follows the structure of the physical contact matrix C^{Base}, with one key modification to the age contact matrix for children. Specifically, we used the same 5×5 regional contact matrix G^{Base} and adjusted the 5×5 age contact matrix A^{Base} to be $A^{\text{OpinionBase}}$. In the opinion sharing age contact matrix, we assume that children's opinions (aged 0 to 17) do not influence the opinions of older individuals. The modified age contact matrix $A^{\text{OpinionBase}}$ is formulated as

$$A_{a_x a_y}^{\text{OpinionBase}} = \begin{cases} 0 & \text{if } a_x \text{is Age 0-17 and } a_y \text{is not Age 0-17} \\ A_{a_x a_y}^{\text{Base}} & \text{otherwise.} \end{cases}$$

We then normalized $A^{\text{OpinionBase}}$ so its sum equals one. Finally, the opinion sharing matrix O^{Base} was calculated using the Kronecker product of G^{Base} and $A^{\text{OpinionBase}}$. Figure B.1.2 provides the contact matrices.

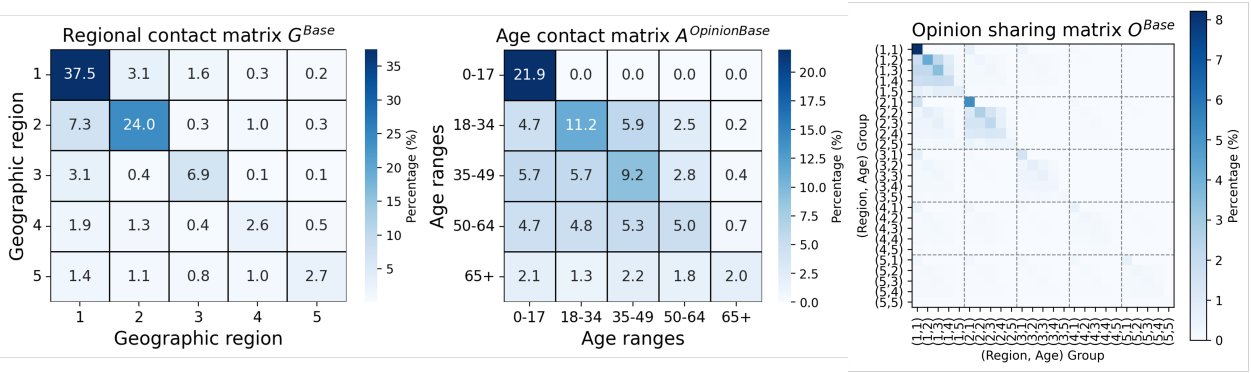


Figure B.1.2: Opinion sharing matrix (O^{Base}) in King County, obtained by taking a Kronecker product of the regional contact matrix G^{Base} and age contact matrix $A^{\text{OpinionBase}}$.

B.1.3 Opinion sharing matrix in online networks O^{Online}

In the sensitivity analysis, we incorporate online social network data to construct the 25×25 online opinion sharing matrix (O^{Online}). We use the Facebook Social Connectedness Index to derive the 5×5 online regional contact matrix G^{Online}. The Facebook index measures the regional connectivity of Facebook users and their friendship networks at the zip code tabulation level [1]. By mapping zip codes to our geographic region level, we calculate the corresponding regional sharing matrix G^{Online}.

Additionally, we obtain the 5×5 online age contact matrix A^{Online} by incorporating social media usage probability by age ranges [3, 14]. Denoting this probability as U_{a_x}, the usage rates are 40%, 84%, 81%, 73%, and 45% for age ranges 0-17, 18-34, 35-49, 50-64, and 65+, respectively. We multiply these probabilities with the previously derived age matrix $A^{\text{OpinionBase}}$. That is,

$$A^{\text{Online}}_{a_x a_y} = U_{a_x} \times A^{\text{OpinionBase}}_{a_x a_y}.$$

We then normalize A^{Online} so that its sum equals one. Finally, the opinion sharing matrix in online O^{Online} was calculated using the Kronecker product of G^{Online} and A^{Online}. Figure B.1.3 provides the contact matrices.

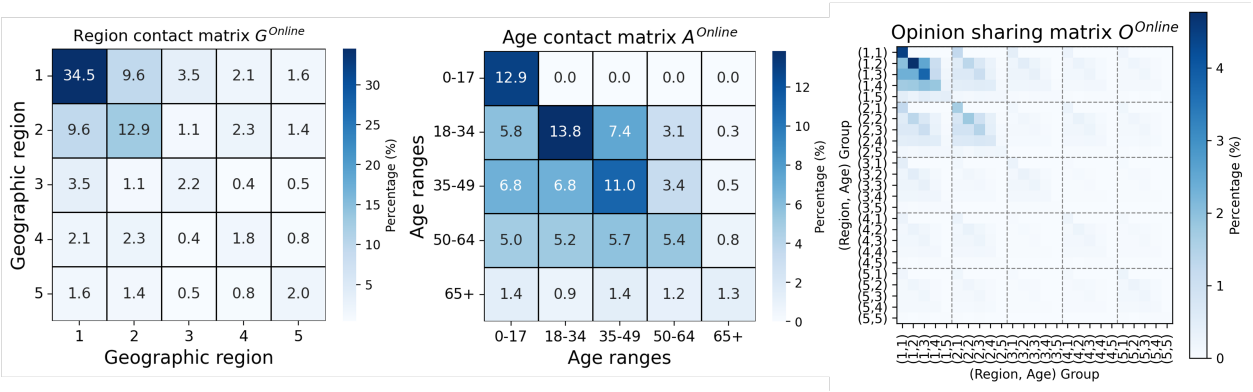


Figure B.1.3: Opinion sharing matrix in online networks (O^{Online}) in King County, obtained by taking a Kronecker product of the regional contact matrix G^{Online} and age contact matrix A^{Online}.

B.2 Initial State Variables

Table B.2.2 presents the initial population and vaccination population rate for each group. The fourth column, initial population $N_i(0)$, was based on 2020 population [13]. The pro-vaccination population rate $P_i(0)/N_i(0)$, was obtained using the bivalent booster vaccination rate for COVID-19 as of January 4, 2023 [5, 22]. Because CDC provides a single rate for the age range 18 to 64 [5], we divided the age range 18 to 64 into three groups (18 to 34, 35 to 49, and 50 to 64) and estimated the vaccination rate for each group by using the relative vaccination rates in WA [22]. We assumed that the vaccination rate only varies by age and not across region groups due to data limits. The anti-vaccination population rate $A_i(0)/N_i(0)$ was calculated by subtracting the pro-vaccination population rate from 1 (i.e., $A_i(0)/N_i(0) = 1 - P_i(0)/N_i(0)$).

Within each of anti- and pro-vaccination groups $(A_i(0), P_i(0))$, individuals can be in one of four disease status (S, I, R, D). We assume that the initial proportion of individuals in each disease compartment is identical across both vaccination groups. Due to data challenges, we calibrated an initial susceptible population rate $s(0)$ that applies to all groups. The initial infectious population rate $i(0)$, as of January 4, 2023, was estimated from summing the reported cases for a ten-day period (January 4 to January 13) and dividing by an estimated infection-detection rate of 0.04 [17]. This resulted in an overall infectious population rate $i(0)$ of 2.68% across all region and age ranges. We model that there is no deceased individuals at the initial setting. The remaining population is initially placed in the recovered compartment. This setup results in the following

initial state variables for each group i,

$$
\begin{aligned}
S_i^A(0)/A_i(0) &= S_i^P(0)/P_i(0) = s(0) \\
I_i^A(0)/A_i(0) &= I_i^P(0)/P_i(0) = i(0) \\
R_i^A(0)/A_i(0) &= R_i^P(0)/P_i(0) = 1 - s(0) - i(0) \\
D_i^A(0)/A_i(0) &= D_i^P(0)/P_i(0) = 0.
\end{aligned}
$$

B.3 Calibration

We calibrate six unknown parameters listed in Table B.3.3. These parameters include four disease-related parameters (α, β, τ_m, $s(0)$) and two opinion-related parameters (o, ϕ). For simplicity, we collectively denote them as $p = (p_1, \ldots, p_6)$. Each parameter has a lower bound and an upper bound, $l_i \leq p_i \leq u_i$ which was derived from literature when available.

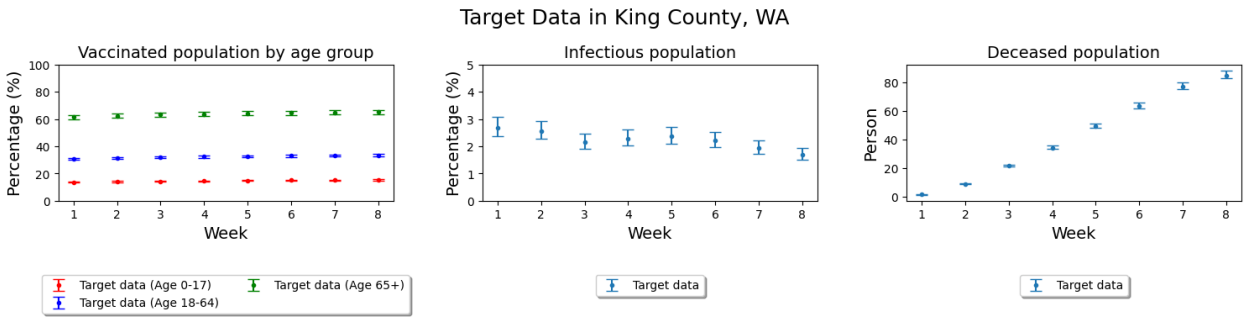


Figure B.3.4: Target data for King County, WA which consists of eight weekly data points on vaccination, infectious, and deceased population from January to February 2023

Figure B.3.4 displays the target data y which consists of eight weekly data points from January to February 2023 and comprises five data types. We use the notation $y_{k,t}$ to represent the target data, where k refers to the type of a health outcome $k \in \{1, \ldots, 5\}$ and t represents each week, starting from the first full week (day 7) up to the eighth week (day 56) (i.e., $t \in \{7, 14, \ldots, 56\}$). Specifically, $k = 1, 2$, and 3 correspond to the percentage of the population that received the bivalent booster vaccination in age ranges 0 to 17, 18 to 64, and over 65 in King County, WA, respectively [5]; $k = 4$ denotes the infectious population percentage in King County, WA, estimated from multiplying reported confirmed cases by infection-detection rate [2, 7, 22]; and $k = 5$ represents the running cumulative number of deaths in King County, WA [22]. We use $g_{k,t}(p)$ to represent the model

health outcomes that correspond to the calibration parameter set p, as below

$$
\begin{aligned}
g_{1,t}(p) &= \frac{\sum_r P_{(r,1)}(t)}{\sum_r N_{(r,1)}(t)}, \quad r \in (1,\ldots,5) \\
g_{2,t}(p) &= \frac{\sum_r \sum_a P_{(r,a)}(t)}{\sum_r \sum_a N_{(r,a)}(t)}, \quad r \in (1,\ldots,5), a \in (2,\ldots,4) \\
g_{3,t}(p) &= \frac{\sum_r P_{(r,5)}(t)}{\sum_r N_{(r,5)}(t)}, \quad r \in (1,\ldots,5) \\
g_{4,t}(p) &= \frac{\sum_r \sum_a I^A_{(r,a)}(t) + I^P_{(r,a)}(t)}{\sum_r \sum_a N_{(r,a)}(t)}, \quad r \in (1,\ldots,5), a \in (1,\ldots,5) \\
g_{5,t}(p) &= \sum_r \sum_a D^A_{(r,a)}(t) + D^P_{(r,a)}(t), \quad r \in (1,\ldots,5), a \in (1,\ldots,5)
\end{aligned}
$$

where groups are represented by (region, age) pair and $t \in \{7, 14, \ldots, 56\}$.

Our goodness-of-fit measure between target data $y_{k,t}$ and model outcomes $g_{k,t}(p)$ is calculated as the sum of mean absolute errors (MAE), which is then normalized by the average of each target data. We use MAE instead of mean squared error (MSE) because our normalized measure ranges between 0 to 1, so the absolute difference provides an estimate of the absolute error, whereas MSE tends to underestimate larger differences in the 0 to 1 range. The calibration optimization problem is formulated as follows:

$$
\begin{aligned}
\min_p \quad & \text{GOF}(y, g(p)) = \sum_{k=1}^{5} \sum_{t \in T_c} \frac{\left| y_{k,t} - g_{k,t}(p) \right|}{\bar{y}_k} \\
\text{s.t.} \quad & l \leq p \leq u \\
& p \in \mathbb{R}^6
\end{aligned}
$$

where $T_c = \{7, 14, \ldots, 56\}$, $\bar{y}_k = (1/|T_c|) \sum_{t \in T_c} y_{k,t}$, and $l, u \in \mathbb{R}^6$.

Finally, we use Latin Hypercube Sampling (LHS) [10] to sample 5,000 parameter sets and choose the one that minimizes the objective function while meet the real-world trends. The best parameter set projected a vaccination uptake of 6.7% without any campaign until end of October, which is similar to observed data during that period (a 6.1% increase from January 1, 2023, to October 31, 2023 [5]). Figure 2 in the main paper shows the calibration result.

Input disease parameters		
Notation	Values	References
C_{ij}	See Figure B.1.1	
β	2.57 per person per day	Calibrated
v_β	0.5	[21]
v_d	0.1	[21]
τ_m	213 days	Calibrated
τ_r	10 days	[4]
τ_d	30 days	[4]
α	0.00013	Calibrated
r_i	$r_{(1,1)} = r_{(2,1)} = r_{(3,1)} = r_{(4,1)} = r_{(5,1)} = 0.16,$ $r_{(1,2)} = r_{(2,2)} = r_{(3,2)} = r_{(4,2)} = r_{(5,2)} = 1.74,$ $r_{(1,3)} = r_{(2,3)} = r_{(3,3)} = r_{(4,3)} = r_{(5,3)} = 7.83,$ $r_{(1,4)} = r_{(2,4)} = r_{(3,4)} = r_{(4,4)} = r_{(5,4)} = 25,$ $r_{(1,5)} = r_{(2,5)} = r_{(3,5)} = r_{(4,5)} = r_{(5,5)} = 100$	[6]
Input opinion parameters		
Notation	Values	References
O_{ij}	See Figure B.1.2	
o	8.75 per person per day	Calibrated
ϕ	2.94%	Calibrated
ρ	0 to 1	Assumed
ν	0 to 0.00003	Assumed
k_R	1/25000 to 1/15000	Assumed
k_E	1/9 to 1/15	Assumed
Vaccination campaign related parameters		
Notation	Values	References
T	365 days	Assumed
n	25 groups	Assumed
C	20,000 individuals	Assumed

Table B.1.1: Parameters values and references

Group	Geographic Region	Age range	Initial Population $N_i(0)$	Initial Vaccination Rate $P_i(0)/N_i(0)$
(1,1)	Seattle	0 to 17	195,404	0.137
(1,2)		18 to 34	341,032	0.243
(1,3)		35 to 49	238,103	0.315
(1,4)		50 to 64	193,590	0.399
(1,5)		65 and over	143,355	0.614
(2,1)	Seattle East	0 to 17	136,108	0.137
(2,2)		18 to 34	139,843	0.243
(2,3)		35 to 49	140,711	0.325
(2,4)		50 to 64	114,934	0.388
(2,5)		65 and over	81,827	0.614
(3,1)	Federal Way	0 to 17	45,210	0.137
(3,2)	-Auburn	18 to 34	44,472	0.243
(3,3)	-Vashon Island	35 to 49	36,373	0.325
(3,4)		50 to 64	38,393	0.388
(3,5)		65 and over	26,364	0.614
(4,1)	Issaquah Plateau	0 to 17	39,394	0.137
(4,2)	-Tahoma	18 to 34	26,527	0.243
(4,3)	-Maple Valley	35 to 49	37,742	0.325
(4,4)		50 to 64	32,360	0.388
(4,5)		65 and over	17,711	0.614
(5,1)	Enumclaw Plateau	0 to 17	31,978	0.137
(5,2)	-Snoqualmie Valley	18 to 34	23,965	0.243
(5,3)		35 to 49	28,111	0.325
(5,4)		50 to 64	26,920	0.388
(5,5)		65 and over	15,075	0.614

Table B.2.2: Initial population and vaccination rate for each group in King County, WA

Parameter	Description	Lower bound	Upper bound	Reference
α	Infection-fatality rate of reference age range (18 to 29 years)	0.0001	0.0002	[15]
β	Transmission rate of infectious individuals	1.5	3.5	Assumed
τ_m	Average number of days from recovery to become susceptible	150	250	[19, 20]
$s(0)$	Proportion of initially susceptible individuals for all groups	0.4	0.8	Assumed
o	Average number of opinion sharing contacts per day when sharing opinion	3	6	[9, 12, 23]
ϕ	Initial proportion of pro-vaccinators that share vaccination opinion	0.01	0.05	[11]

Table B.3.3: Calibration parameters and their upper and lower bounds

REFERENCES FOR APPENDIX B

1. Bailey M, Cao R, Kuchler T, Stroebel J, and Wong A. Social Connectedness: Measurement, Determinants, and Effects. Journal of Economic Perspectives 2018 Aug; 32:259–80. DOI: 10.1257/jep.32.3.259

2. Barber RM, Sorensen RJD, Pigott DM, Bisignano C, Carter A, Amlag JO, et al. Estimating global, regional, and national daily and cumulative infections with SARS-CoV-2 through Nov 14, 2021: a statistical analysis. The Lancet 2022; 399:2351–80. DOI: 10.1016/S0140-6736(22)00484-6

3. Canales K. 40% of kids under 13 already use Instagram and some are experiencing abuse and sexual solicitation, a report finds, as the tech giant considers building an Instagram app for kids. [Online; accessed 6-March-2024]. 2021. Available from: https://www.businessinsider.com/kids-under-13-use-facebook-instagram-2021-5

4. Centers for Disease Control and Prevention. COVID-19 Pandemic Planning Scenarios. [Online; accessed 15-April-2023]. 2020. Available from: https://www.cdc.gov/coronavirus/2019-ncov/hcp/planning-scenarios.html

5. Centers for Disease Control and Prevention. COVID-19 vaccinations in the United States, County. [Online; accessed 24-March-2023]. 2023. Available from: https://data.cdc.gov/Vaccinations/COVID-19-Vaccinations-in-the-United-States-County/8xkx-amqh

6. Centers for Disease Control and Prevention. Risk for COVID-19 infection, hospitalization, and death by age group. [Online; accessed 24-March-2023]. 2022. Available from: https://archive.cdc.gov/www_cdc_gov/coronavirus/2019-ncov/covid-data/investigations-discovery/hospitalization-death-by-age.html

7. Institute for Health Metrics and Evaluation. COVID-19 Projections. [Online; accessed 10-January-2023]. 2020. Available from: https://covid19.healthdata.org/projections

8. Kang Y, Gao S, Liang Y, Li M, Rao J, and Kruse J. Multiscale dynamic human mobility flow dataset in the US during the COVID-19 epidemic. Scientific data 2020; 7:390

9. Macdonald B, Luo M, and Hülür G. Daily social interactions and well-being in older adults: The role of interaction modality. Journal of Social and Personal Relationships 2021; 38:3566–89. DOI: 10.1177/02654075211052536

10. McKay MD, Beckman RJ, and Conover WJ. A Comparison of Three Methods for Selecting Values of Input Variables in the Analysis of Output from a Computer Code. Technometrics 1979; 21:239–45

11. Morning Consult & The Hollywood Reporter. Share of adults in the United States who have shared an opinion piece they agreed with on social media as of November 2018, by age group. [Online; accessed 16-January-2023]. 2018. Available from: https://www.statista.com/statistics/990249/agreed-with-opinion-piece-shared-social-media-usa/

12. Nguyen L, Berg P van den, Kemperman A, and Mohammadi M. Where do People Interact in High-Rise Apartment Buildings? Exploring the Influence of Personal and Neighborhood Characteristics. International Journal of Environmental Research and Public Health 2020; 17

13. Office of Financial Management. Census geographic files - Census 2020. [Online; accessed 24-March-2023]. 2023. Available from: https://ofm.wa.gov/washington-data-research/population-demographics/gis-data/census-geographic-files

14. Pew Research Center. Social Media Fact Sheet. [Online; accessed 6-March-2024]. 2024. Available from: https://www.pewresearch.org/internet/fact-sheet/social-media/?tabId=tab-d102dcb7-e8a1-42cd-a04e-ee442f81505a

15. Pezzullo AM, Axfors C, Contopoulos-Ioannidis DG, Apostolatos A, and Ioannidis JP. Age-stratified infection fatality rate of COVID-19 in the non-elderly population. Environmental Research 2023; 216:114655. DOI: https://doi.org/10.1016/j.envres.2022.114655

16. Prem K, Cook AR, and Jit M. Projecting social contact matrices in 152 countries using contact surveys and demographic data. PLoS computational biology 2017; 13:e1005697

17. Qasmieh SA, Robertson MM, Teasdale CA, Kulkarni SG, Jones HE, McNairy M, et al. The prevalence of SARS-CoV-2 infection and long COVID in US adults during the BA. 4/BA. 5 surge, June–July 2022. Preventive Medicine 2023; 169:107461

18. SafeGraph. The impact of coronavirus (COVID-19) on foot traffic. [Online; accessed 24-March-2023]. 2020. Available from: https://www.safegraph.com/dashboard/covid19-commerce-patterns

19. Stein C, Nassereldine H, Sorensen RJD, Amlag JO, Bisignano C, Byrne S, et al. Past SARS-CoV-2 infection protection against re-infection: a systematic review and meta-analysis. The Lancet 2023; 401:833–42. DOI: 10.1016/S0140-6736(22)02465-5

20. Swartz MD, DeSantis SM, Yaseen A, Brito FA, Valerio-Shewmaker MA, Messiah SE, et al. Antibody Duration After Infection From SARS-CoV-2 in the Texas Coronavirus Antibody Response Survey. The Journal of Infectious Diseases 2022 May; 227:193–201. DOI: 10.1093/infdis/jiac167

21. UK Health Security Agency. COVID-19 vaccine surveillance report - Week 24. Web Page. [Online; accessed 14-September-2022]. 2022. Available from: https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1083443/Vaccine-surveillance-report-week-24.pdf

22. Washington State, Department of Health. COVID-19 data dashboard. [Online; accessed 24-March-2023]. 2023. Available from: https://doh.wa.gov/emergencies/covid-19/data-dashboard

23. Zhaoyang R, Sliwinski MJ, Martire LM, and Smyth JM. Age differences in adults' daily social interactions: An ecological momentary assessment study. 2018. DOI: 10.1037/pag0000242