AN INTEGRATED EPIDEMIOLOGICAL SUPPLY CHAIN MANAGEMENT APPROACH FOR THE COVID-19 VACCINE DISTRIBUTION AND ALLOCATION PROBLEM

by

Jacob David Locke

Submitted in partial fulfillment of the requirements
for the degree of Master of Applied Science

at

Dalhousie University
Halifax, Nova Scotia
August 2023

Dalhousie University is located in Mi'kma'ki, the ancestral and unceeded territory of the Mi'kmaq. We are all Treaty people.

Table of Contents

List of Tables

List of Figures

Abstract

This thesis proposes an integrated model for vaccine distribution and allocation that combines an SEIR compartmental model with a transshipment model. Given the NP-hardness of the problem, eight solution methods are explored and evaluated in terms of solution quality and run time. The best-performing method was found to be a greedy marginal benefit heuristic combined with a genetic algorithm while a logic-based Benders Decomposition approach provided provably optimal solutions to homogeneous vaccine allocation problems, albeit in prohibitively long solve times. Applied to a case study based on data from Ontario during the COVID-19 pandemic, we identify the best practices for allocating and distributing a limited supply of vaccines to minimize both total cases and deaths. The best identified vaccination strategies were able to reduce the total number of cases by 25% compared to a Pro-rata allocation of vaccines while saving $6,800,000 in vaccine acquisition and transportation costs.

List of Abbreviations Used

BD Benders Decomposition.

COVID-19 Coronavirus disease 2019.

DC Distribution center.

EPI Expanded Program on Immunization.

GA Genetic Algorithm.

GMB Greedy Marginal Benefit.

MC Mountain Climbing.

NI Non-Intervention.

PHU Public Health Unit.

PR Pro-rata.

SEIR Susceptible-Exposed-Infected-Removed.

SIR Susceptible-Infected-Removed.

SVEIR Susceptible-Vaccinated-Exposed-Infected-Removed.

WHO World Health Organization.

YLL Years of Life Lost.

Acknowledgements

I would like to thank my supervisors Dr. Ahmed Saif and Dr. Bahareh Mansouri, as well as Dr. Majid Taghavi for their continuous support and guidance throughout this journey in completing my Master's thesis. Their encouragement and constructive feedback were instrumental in shaping the direction of this research and refining the quality of the final work and without their mentor ship I would never have been able to reach this point.

I would also like to thank my friends and family for their support, understanding, and encouragement throughout the process. Without them, I would never have been able to reach this point.

Chapter 1

Introduction

In early 2020, the world experienced a healthcare emergency with the emergence of COVID-19 and its subsequent spread around the globe. This sparked an extraordinary effort by the scientific community to develop a vaccine, culminating in not just one, but several vaccines being developed and approved for public use in less than a year. But that was only the beginning of the end, as the task then became how to best allocate and distribute a limited supply of vaccines as governments around the world competed for these valuable resources [10]. It was important to make the best use of limited vaccines to maximize the societal benefit by reducing and controlling the extent of the outbreak while accounting for the logistical limitations and minimizing the economic burden of distributing large quantities of vaccines. Policymakers had to decide who should have priority and what should guide these decisions. How would these vaccines be distributed to where they were needed and was the chosen strategy even going to be feasible?

1.1 Background

Vaccines are a vital tool in public health and one of the most cost-effective methods of curbing the spread and impact of a disease. Used preventively, such as in childhood and seasonal vaccination campaigns, they help prevent outbreaks from occurring and have successfully eradicated or nearly eradicated diseases such as smallpox, polio, and measles. Used reactively in the event of an outbreak, they help prevent the worse symptoms of infectious diseases, while also lowering transmission rates, both of which help save lives and prevent outbreaks from overwhelming healthcare systems [4, 19]. However, without a proper logistics network behind them, vaccines would be far less effective. Vaccines are a unique product. They are subject to strict health and safety regulations not only in their production, but also in their packaging, transportation, and storage. Additionally, unlike some other medical products, most vaccines are

not stable at room temperature, and as such must be stored and transported at refrigerator, freezer, or even deep-freeze temperatures. This makes the vaccine supply chain a cold chain, requiring specific infrastructure that leads to increased costs and energy consumption. In developing countries, cold chain resources are often limited, with no way for some communities to safely store vaccines for prolonged periods. Even in ideal storage conditions, vaccines cannot be stored indefinitely, meaning stockpiling vaccines in preparation for an outbreak inevitably leads to at least some wasted stock. [3, 26].

As noted by Duijzer, Jaarsveld, and Dekker [19], in the event of a sudden outbreak, the available doses of vaccine are often insufficient to vaccinate the entire population and it can take a long time for production to meet demand. On top of the distribution problem, this situation creates an allocation problem in deciding who should be vaccinated. Populations can be divided into high- and low-risk and also high- and low- transmission rates. In the event of a multi-region outbreak, geographic considerations have to be accounted for as well. Careful analysis of the situation is required to decide who should have priority. In such situations, the use of epidemiological models can be useful to predict the course of the disease, forecast demand, and make informed decisions about where vaccines should be allocated to control the epidemic.

There are two primary methods of modeling an epidemic, agent-based simulation models, and compartmental models, of which compartmental models are the more popular option due to their simplicity, flexibility, and reduced computational power required compared to agent-based simulations [15]. First developed by Kermack and Mckendrick in 1927 [33], compartmental epidemiological models work by dividing the population into distinct groups based on whether they are susceptible to catching the disease, are currently infected, or have recovered or died from the disease. Differential equations are used to determine the rate of transfer between these compartments with respect to time, which can be approximated using difference equations when discrete-time models are used. One such compartmental model that serves as the basis for the majority of other models is the Susceptible-Infected-Removed (SIR) model [49, 57, 59]. Common variants of SIR are SEIR models [21, 47], which add a compartment to represent those who have been exposed to the disease but are not yet infectious,

and SVEIR models [5, 44] which add different compartments to represent vaccinated individuals. Even recent models, such as the DELPHI model developed specifically to model the COVID-19 pandemic, are based around the SIR compartmental model [9, 35]. However, the dynamics of disease spread mean that the number of new infections is proportional to the product of the remaining susceptible population and the current number of infected individuals, leading to nonlinear dynamic models. Hence, to minimize the number of new infections, one has to solve a non-convex problem that is known to be NP-hard [43, 55]. This means that while simulating the course of a disease is relatively easy, determining an optimal course of action can prove exceptionally difficult.

Before the outbreak of COVID-19, many researchers had studied the problem of the vaccine supply chain and identifying optimal vaccination strategies, but the pandemic only served to further highlight the importance of effective vaccination campaigns and catalyzed an acceleration in the number of new studies on the topic as well as a change in the research agendas [10]. For instance, studies on vaccine allocation for seasonal influenza often focused on identifying the minimum fraction of the population required to be vaccinated to get an outbreak under control [21, 49]. The assumption was that vaccine supply would not be a limiting factor and that the disease could be stopped before it got out of control. By the time vaccines for COVID-19 were being made available, the disease was already well-established in the population and there was only a limited supply of vaccines. Containing the outbreak was infeasible under such conditions. Instead, the focus of COVID-19 vaccine allocation studies is usually on the allocation of a limited supply of vaccines to mitigate infections, deaths, YLL, etc. [44, 63]. Similarly, previous distribution studies were unprepared for the scale of the COVID-19 mass vaccination campaign, necessitating re-examining how vaccine supply chain and distribution networks should be structured [40]. Many existing studies do not consider both the allocation and distribution aspects of vaccination, instead focusing on one or the other. Distribution studies usually have a predetermined demand and focus on costs and fulfillment, without seeking to control the disease or modeling its impact on the course of the outbreak [18, 22, 36]. On the other hand, many vaccine allocation studies have not accounted for the supply chain limitations surrounding implementing their proposed

strategies, in reality, [21, 41, 49].

Likewise, vaccine allocation studies often do not consider allocating individual vaccines directly, but instead, find what fraction of the population should be vaccinated. They assume that all vaccination can be done all at once (i.e., in a single period) [21, 41, 49], rather than over time (multiple periods) [9, 44, 59]. Vaccine allocation studies commonly use homogeneous epidemiological models [35, 57] in place of heterogeneous ones [9, 44], meaning groups and locations are kept separated with no interaction, with each population using its own independent epidemiological model. These assumptions and alterations simplify the vaccine allocation problem and make it easier to solve, but also fundamentally change the nature of the system being modeled.

1.2 Summary of Work

This thesis presents a new integrated model that combines a heterogeneous epidemiological model for multi-period vaccine allocation and a vaccine supply chain model, and explores several solution methods for the vaccine allocation problem, comparing their solution quality and resolution times. The epidemiological model used in this thesis is a modified version of a traditional SEIR model and is used to capture the dynamics of the disease spread and how different vaccination strategies impact them. The vaccine supply chain model follows the formulation of a standard network problem, with a pool of vaccines made available for distribution to a set of population centers during each period through a set of distribution centers (DC). These two models work in tandem to identify vaccine allocation strategies that minimize the epidemiological impact of an outbreak both within and across population centers while balancing against the economic cost and logistical limitation on the system when trying to fulfill the selected vaccination strategy.

The vaccine allocation formulation presented in this thesis is highly complex, as it is non-linear, non-convex, and combinatorial, resulting in a large-scale mixed-integer global optimization problem. As such, it is extremely difficult to solve to the proven optimality. To the best of our knowledge, no previous study has been able to present a method to solve the vaccine allocation problem to the guaranteed optimal within a reasonable time frame without making gross simplifications such as linearizing

the model [12, 24], or significantly limiting the problem size [21, 59]. Therefore, we explore eight different solution methods, explain how they function, and present the advantages and disadvantages of each method. It was found that four of the explored methods were promising, being able to produce near-optimal solutions in a reasonable time (i.e., no more than 24 hours), for problems of a reasonable size (i.e.,10 or more distinct populations, 5 or more risk groups, and at least 20 time periods). These methods included three heuristic methods: a Mountain Climbing (MC), a Greedy Marginal Benefit (GMB), and a Genetic Algorithm (GA). Alongside these heuristic methods is a novel method employing logic-based Benders Decomposition (BD), which guarantees optimality, though its solution time in all but the simplest test scenarios was found to be prohibitively long. We hope that by presenting the full scope of the explored methods (promising or not), the community might be informed of how a broad range of methods perform and potentially allow future work to build off our own to develop further methods to solve the problem of vaccine allocation and distribution.

The solution methods are tested under a variety of scenarios to evaluate their performance in terms of both the solution quality (their epidemic impact and monetary cost), and run time. Performance is measured relative to the baseline scenario of Non-Intervention (NI) and against simple rule-of-thumb strategies such as Pro-rata (PR) (proportional to population size) allocation. From this round of experiments, we identify that the best-performing solution method of those explored in terms of solution quality was the GMB combined with the GA, which can produce a near-optimal solution in just over 24 hours for a problem of reasonable size (e.g., 34 regions, 7 groups, 20 time periods), with potential for further improvements in both speed and quality. However, it might be better to simply use the GMB alone, as the difference in solution quality could be considered insignificant for the additional run time added by the GA. The fastest solution method that could still generate a reasonably good solution was found to be the MC, though for larger problems there was around an 11% gap between its solution and that of the GMB. However, there is potential to improve its quality in future research. In addition to these heuristic methods, we show that our BD method can solve a problem with a single risk group but multiple regions to the guaranteed optimal, though the time it takes to solve

all but the smallest problems is prohibitively long. However, we do identify several potential directions to explore that might lead to improvements for this method.

Finally, the model and solution methods are used to explore a case study based on the Canadian province of Ontario during the COVID-19 pandemic. Using the best solution methods, approximately 683,500 cases were prevented over 20 weeks when compared to the NI policy, while 419,700 cases were prevented when compared to the PR allocation policy, with cost savings of $6,798,000. When the objective was to minimize deaths rather than cases, 1,260 deaths were prevented compared to the NI policy, and 350 deaths were prevented compared to the PR allocation policy, with cost savings of $25,612,000. As the model is multi-objective, balancing epidemic impact (cases, deaths, etc.) against monetary costs, these values represent a conversion factor between the two objectives of decision makers being willing to spend $1,000 per prevented case, and $50,000 per prevented death, respectively.

Additionally, from these case study results, insights can be derived that might benefit decision-makers such as governing bodies or health authorities in deciding how to best utilize limited resources during an epidemic situation, even without directly utilizing the solution methods. Most notably, when preventing infections, the optimal strategy is to focus on young children and teenagers, as they have many more contacts per week than any other groups [11], thus acting as a super vector for spreading the disease to others. However, if looking to prevent deaths, a combination of vaccinating older age brackets; as they are the most vulnerable, and vaccinating the young super spreaders, thus slowing the spread through the rest of the population, is a better strategy. In both cases, when supplies are limited, focusing on just a few areas rather than spreading resources equally across the entire region was found to be the optimal choice. It was found that the monetary amount that the decision-makers are willing to spend to prevent additional infections or deaths impacts both the number of vaccines utilized and the locations where they are sent. As cost restrictions tighten, the number of allocated vaccines will begin to drop rapidly and large increases in cases and deaths are observed, especially for remote areas that are further from DCs and have rather high shipping costs.

1.3 Contributions

In summary, this thesis makes four contributions. 1) From a modeling perspective, it proposes a novel integrated model that combines features of SEIR epidemiological and supply chain models to formulate a multi-period vaccine allocation strategy. 2) It explores eight solution methods for solving the problem of vaccine allocation and distribution, including three promising heuristic methods along with an exact novel approach using logic-based BD with a guarantee of optimality. 3) It applies the model and solution methods to a case study with real data derived from the province of Ontario's COVID-19 statistics. 4) Finally, it provides practical insights into where to best allocate vaccines to make the most of limited resources.

1.4 Thesis Outline

The rest of this thesis is organized as follows: Chapter 2 covers a literature review of existing research on vaccine allocation. Chapter 3 describes the problem structure and the model formulation. Chapter 4 presents the solution methods. Chapter 5 outlines the numerical results and the evaluation of the different solution methods. Chapter 6 applies the promising solution methods to the case study, outlines the results, and presents what insights were gained from their application. Chapter 7 provides the thesis conclusions and suggests promising directions for future research.

Chapter 2

Literature Review

Since the mid-1900s, Operations Research has been used as a standard tool for managing healthcare logistics, specifically resource planning, scheduling, and allocation. A review of existing literature finds that there has been a steady interest in the topic of modeling the spread of epidemic disease and developing strategies to control outbreaks using various measures. The emergence of COVID-19 onto the world stage in early 2020 highlighted the importance of making the most of limited resources and having reliable and fast supply chains to produce, distribute, and administer vaccines. As such, there has been an increase in the number of papers studying vaccine supply chains and how best to use vaccines to combat the spread of disease. This section reviews those studies, focusing on those most related to the current study.

There have been several survey papers that analyzed vaccine supply chain literature until 2020. Lemmens et al. [34] reviewed supply chain network design models with a focus on vaccine supply chains. They discussed the strategic, tactical, and operational decisions a supply chain has to make, and categorized the reviewed references based on network characteristics, performance measures, and solution methodology used. While most studies are focused on minimizing the cost of supply chain network design, it has been argued that there is value in designing an equitable vaccine supply chain and/or ways to measure the humanitarian aspect of the vaccine supply chain. Duijzer, Jaarsveld, and Dekker [19] characterized vaccine supply chain studies by product, production, allocation, and distribution models, and identified the cost-effectiveness of the vaccine supply chain, preparation for outbreaks, and preparation for bio-terror attacks to be the main challenges that vaccine supply chains are facing. De Boeck, Decouttere, and Vandaele [17] characterized vaccine supply chains by studying sourcing, storage, transportation, and administration of vaccines and dividing the decisions into strategic, tactical, and operational levels. They reported that simulation is the most widely-used methodology, followed

by optimization techniques, and then analytical methods. They concluded that no study has considered all four characteristics, from sourcing to administration, of the vaccine supply chain under uncertainty.

A recent review by Blasiolo et al [10] focused on vaccine distribution and allocation studies in the context of how they were or could be applied to COVID-19. They identified several issues that made the equitable allocation of vaccines during the pandemic difficult, namely 1) the unprecedented scale resulting in fierce competition for access to the limited supply of vaccines, 2) the urgency in development which led to increased hesitancy, 3) production and distribution challenges from the huge demand and complexity of the supply chain, 4) the use of mRNA which induced additional pressure from scarcity of materials and skilled labor, and 5) shortage gaming and global politics resulting in high-income countries taking more of the limited supply of vaccines than they needed. In the context of vaccine allocation, they highlighted the need to develop models that can further incorporate equity, such as including disparities in accessing healthcare, equity in allocation, and vaccine hesitancy. They highlight a need for studies that incorporate prioritizing allocation among patients with comorbidity and high-risk workers, capturing interaction amongst individuals according to population density and age, capturing the effect of second and third doses and different vaccine types, and hesitancy towards vaccination.

In this thesis, the review focuses primarily on studies on vaccine supply chains, distribution, and allocation. The literature is categorized based on its primary focus and further subdivided based on methodology and objectives. In section 2.1 we discuss studies that focus on the vaccine supply chain, which generally covers the period from manufacturer to administration. These studies generally discuss questions such as where to locate facilities, capacity considerations, and how to best move vaccines through a distribution network. In general, these studies either have fixed or stochastic demand for vaccines that do not account for disease patterns or impact. They mainly aim to minimize the economic, environmental, and/or societal costs incurred in the distribution of vaccines to meet a given demand, or improve the availability of vaccines by making the most of limited resources [13, 29, 28, 18, 36, 40, 64, 58, 66, 22, 65, 56].

Sections 2.2 and 2.3 cover studies that use epidemiological modeling to predict and

control the spread of the disease through intervention and control measures. These studies can be further categorized with respect to methodology (such as simulation, optimization techniques, or analytical methods), or defined target and/or impact metric, (such as minimizing costs, total exposures, deaths, YLL, outbreak duration). Additionally, several reviewed studies also included different metrics of equity in their decision-making, such as enforcing that a minimum fraction of a group must be vaccinated, or minimizing the maximum difference between the vaccinated fraction of each group [5, 20, 21, 38, 41, 47, 49, 59, 62, 63].

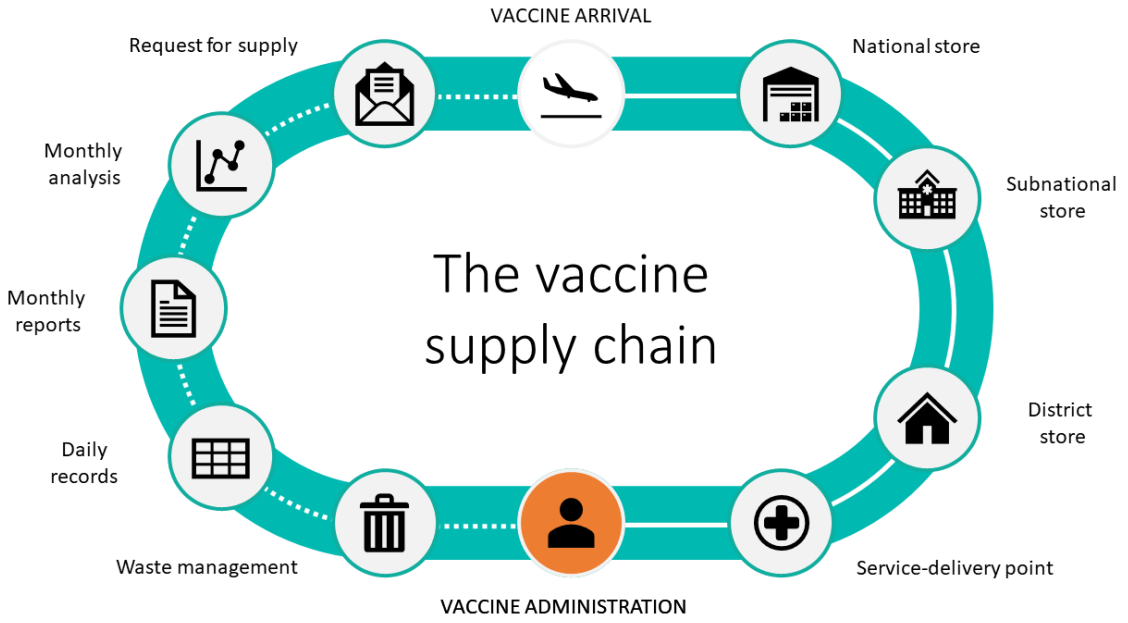
Section 2.4 focuses on studies that combine vaccine allocation, distribution, and epidemic modeling for the purpose of the allocation of resources to control the impact of disease. Similar to the previous categories, these papers are further classified by methodology and their primary objective [9, 40, 44, 46, 58, 66]. This thesis is most relevant to studies reviewed in this category.

2.1 Vaccine Supply Chain and Distribution

A variety of vaccine supply chain issues have been explored by operational researchers. Several tools including stochastic programming, dynamic programming, and network design modeling have been used to model and solve these problems, usually for the purposes of reducing costs, both economical and environmental, or fixing inefficiencies in the distribution network in order to increase throughput and access to vaccines. Occasionally, there are studies that are primarily focused on promoting equitable access to vaccines.

In 1974, the World Health Organization (WHO) launched a program called the Expanded Program on Immunization (EPI) that aimed to improve childhood access to various vaccines in some of the poorest countries around the globe [4, 13]. The program is a collaboration between WHO, UNICEF, the World Bank, various public health departments and charities, and vaccine manufacturers, and aims to help decision-makers plan out and coordinate various aspects of national vaccination strategies. This includes a suggested structure (see Figure 2.1) for the vaccine supply chain of a country that covers the point from which vaccines enter the country to the point they are distributed to local clinics. The suggested structure is a multi-tier network with vaccines flowing down from national stockpiles through several layers

Figure 2.1: Supply chain and logistics and IA2030. Source: World Health Organization [54]



of regional warehouses before ultimately being delivered to clinics where they will be administered. This network structure allows for an organized yet still flexible top-down diffusion of vaccines.

Even though the WHO's EPI provides guidelines for how to develop and structure a national vaccine distribution network, there have been numerous researchers over the years who have questioned whether the suggested framework is the best for all scenarios and where improvements might be made. Among them are Chen et al. [13], who formulated a model that sought to address several of the issues they identified with the WHO's EPI recommendation for the design of a national vaccine distribution network, and presented a framework for expanding on the recommendations from the WHO to tailor it to a specific country's needs. The objective of the model was to maximize the number of fully immunized children by deciding how many vaccines of differing types to store in different facilities across the country at any given time, where to send vaccines to and from, how much basic, refrigerated, and freezer capacity to build into a location, and how much transportation capacity to build into connections between locations. The model used vaccine packaging

volume to measure its capacity usage and also accounted for product loss in storage and transport, with the ability to have different values for each location and vaccine type across different time periods. Applied to a case study in Niger, they evaluated the impact changes such as removing regional-level hubs, changing vial sizes, and introducing new vaccines would have on the national vaccine distribution network.

A similar study by Hovav and Tsadikovich [29] used an optimization model to minimize the setup, distribution, and administrative cost across the entire vaccine supply chain from manufacturer to clinic for seasonal influenza vaccines, with unmet demand for vaccines incurring additional costs on the clinic side through future outpatient and hospitalization services. Their model was able to improve the performance of the real strategies used by Israeli HCO CLALIT Health Services experts during an outbreak in 2013 by approximately 12%. They identified a behavior they called "planned shortages", where in the last few weeks of the season, it was more profitable to stop ordering and administering vaccines and simply incur the additional treatment costs. Building on that study, Hovav and Herbon [28] presented a detailed logistics model to determine the number of shipments and quantity of seasonal influenza vaccines in each periodic shipment that is delivered from manufacturers to DCs, from DC to clinics, and from clinics to sub-groups. The objective of the model was to minimize costs across the entire supply chain for a healthcare organization planning a large-scale vaccination program, including service costs to administer medical care, cost of bad reputation from unmet demand, and the cost saving of health benefits for the customer converted into monetary terms. They found that their approach was particularly useful in situations where resources were limited, and they were able to significantly reduce costs incurred from backlogs, hospitalizations, and outpatient visits.

De Carvalho, Ribeiro, and Barbosa-Póvoa [18] outlined a model for designing and planning sustainable vaccine supply chains by maximizing economic (NPV) and social benefit (GDP) while minimizing environmental impact (ReCiPe). The supply chain was comprised of factories, warehouses, and markets. Vaccines could flow from factories to warehouses or directly to markets. On the facility side, the model could determine where to locate facilities, what their capacities should be, what technologies they should use, and their optimal inventory levels. On the distribution side, it

would decide how much to transport between locations and how many vehicles and trips are required to transport them. The study accounts for numerous auxiliary variables such as net present value, cash flow, net earnings, depreciation capital, and fixed capital investment. Through the application of the model to the creation of a European vaccine supply chain under five different scenarios, they identified the trade-offs that occur when trying to balance economic, social, and environmental concerns. They demonstrated that significant environmental and social improvements can be achieved with minimal compromise in economic return.

Saif and Elhedhli [61] presented a mixed-integer concave model for the purposes of minimizing capacity, transportation, and inventory costs along with greenhouse gas emissions across a cold supply chain. Utilizing a novel application of Lagrangian decomposition to solve the model, they applied their model to several different applications of a cold chain, including a case for distributing vaccines in Ontario. Examining the trade-off that occurs between balancing total cost and total greenhouse gas emissions, they found that as the model prioritized emissions more heavily the transportation costs would dominate over inventory costs and the optimal solution was to open more DCs in order to minimize travel costs.

Expanding on the work by Chen et al. [13], Lim, Norman, and Rajgopal [36] presented a mixed-integer programming model to reorganize vaccine distribution networks in low- and middle-income countries to minimize annual costs while continuing to meet the requirements outlined by the WHO-EPI. By restructuring a subsection of the national distribution network away from its normal rigid hierarchy, optimizing inventory and delivery rates, and optimizing the use of resources such as different types of transportation and vaccine storage devices, they were able to significantly reduce the transportation, storage, and facility costs of the network. Solving this problem to optimality would sometimes take in excess of 196 hours. Therefore, alongside the model, they presented a heuristic to solve larger vaccine distribution network problems to near-optimally in a reasonable time. They concluded that using their model to adapt the vaccine distribution chain to the specific characteristic of a country provided annual savings ranging from just under 10% to just over 30%.

The pandemic highlighted a fundamental need to re-examine how vaccine supply chains should be structured. Manupati et al. [40] argue that most studies that

explored optimizing vaccine distribution prior to the pandemic were not suitable for the magnitude of the distribution efforts required to vaccinate entire populations of people. To remedy this, they formulated their own multi-echelon dynamic cold chain distribution model, specifically catered to mass vaccination drives. They used decision tree analysis to develop different vaccination strategies and then used the synthetic control method to identify the expected reduction in the number of cases and deaths from implementing the strategy. A stochastic mixed-integer programming problem is then used to determine the location and number of cold storage facilities, their required inventory levels, and the allocation of vaccines to clinics, which minimizes both transportation time and cost for a given strategy. A similar paper by Sripada et al. [64] presents a decision support framework for designing a multi-tier vaccine cold chain network along with two optimization formulations to minimize inventory, ordering, transportation, personnel, and shortage costs across a single- and multi-vaccine distribution network, respectively. The framework informs those in lower tiers about which higher tier facilities they should order from, the number of vaccines they should order in each period, the number of vehicles required to transport said vaccines, the inventory levels of each facility in the network for each period, the number of staff required to run each facility, and the number of vaccines to administer to a subgroup of the population in a given time period. Applying this framework to the Indian state of Bihar, they were able to make several observations about how the optimal strategy would change based on different model parameters, such as what impact the volume of packaging would have on the optimal ordering, inventory levels, shortages, and staffing requirements.

Concerned with shortages of influenza vaccines in developing countries amidst the COVID-19 pandemic, Rastegar et al. [58] proposed a single-product, multi-period distribution model for the equitable distribution of vaccines in situations where demand outstrips supply. The model maximizes the minimum delivery-to-demand ratio amongst provinces in a country for a given budget by deciding which DCs to open and how to best allocate vaccines to different risk groups, who each had their own demand for vaccines. They grouped the population into eight risk groups that, rather than primarily focusing on age, consisted of various different high-risk groups such as infants and toddlers, pregnant women, healthcare providers, adults

over 65, and people with pre-existing conditions, with everyone else grouped into the "other" category. The model was then applied to an Iranian case study to suggest the equitable and fair distribution of the influenza vaccine for the fall and winter flu seasons. An extension of this study by Tavana et al. [66] incorporated cold, very cold, and ultra-cold supply chain constraints into decision-making, with the assumption that the distribution process starts when an order is placed with a manufacturer. They applied their model to a case study in India. These models are semi-unique in that they are almost entirely focused on maximizing equity for a given budget, whereas others often balance minimizing economic costs against equity-enforcing constraints.

Fadaki et al.[22] presented a transshipment model for the distribution and allocation of vaccines between a single Distribution center (DC) and a network of medical centers while minimizing exposure risk to the unvaccinated population. Limited vaccines were made available each period, stored in the DC, and distributed out to medical centers. Unvaccinated individuals booked an appointment at a medical center and were prioritized for vaccination based on their susceptibility and exposure risk. Each medical center had only a certain capacity to administer vaccines each period, with unmet demand being carried over to the next period. What distinguished this model from others was that there was a single DC but it allowed for the transshipment of vaccines between medical centers to cover demand along with the return of unused vaccines to the DC. Shipping to and from the DC was unlimited, while there was limited capacity to transport vaccines between medical centers each day. This essentially allowed the model to turn certain medical facilities into hubs that then distributed vaccines to other centers. For a case study in Australia, through their multi-period model, they were able to shorten the administering period and reduce total residual risk significantly compared to the single-period model.

These papers primarily focused on high-level strategic decisions, such as facility location, capacity, inventory levels, and the number of vaccines to send from one location to another, though some did factor vehicle, administrative, and or worker costs and constraints into their models [64]. However, all of these papers considered the delivery of vaccines through the network to be a direct distribution. Vehicles only travel between two points to pick up and drop off their load of vaccines. A

different approach is to examine the operational side and how to actually deliver vaccines on a daily or weekly schedule using a fleet of vehicles. One such study by Sun, Andoh, and Yu [65] used the anyLogistix simulation package to analyze the performance of the distribution of COVID-19 vaccines through the cold chain in a real-world case study in the Oslo area and Viken county in Norway. A fleet of vehicles delivered the daily demand for vaccines to various locations around Oslo and Viken county starting from a regional warehouse. By varying metrics such as fleet size and composition (including the consideration of using unmanned aerial vehicles) they developed, analyzed, and compared twelve different scenarios, and evaluated their performance based on metrics such as transportation costs and emissions, vehicle utilization, service level, and lead time.

All papers discussed thus far have only considered the vaccine supply chain from at most the point where vaccines leave the manufacturer to the point they are administered at a clinic. In contrast, Pishvaee, Razmi, and Torabi [56] proposed a stochastic multi-objective programming model to assist in designing an entire sustainable supply chain for the sourcing, production, distribution, and disposal of medical needles, down to individual components. The model balanced the competing objectives of minimizing total costs, minimizing environmental impact, and maximizing the societal impact in the form of created jobs, local development, and lowered consumer and worker risk.

2.2 Epidemiological Models

Some of the first instances of mathematical modeling being applied to analyzing and predicting the spread of disease occurred in the early twentieth century, with some of the most influential early work being done by Kermack and Mckendrick in 1927 who developed the first SIR (Susceptible-Infected-Removed) compartmental models for simulating the spread of disease [33]. Compartmental epidemiological models work by dividing the population into discrete groups (compartments) based on whether they are susceptible to catching the disease, are currently infected, or have recovered or been killed by the disease. Differential equations are used to determine the rate of transfer between these compartments with respect to time. These first compartmental models were still relatively simple and used a single homogeneous population,

rather than subdividing the population into different regions and risk classes, as would become common practice later. The primary focus of the initial study was to investigate the relationships between population density and transmission rate to the final size and duration of an outbreak. Through this work, it was identified that for a given transmission rate, if the population density of susceptible people does not pass a certain threshold, the disease will quickly die out, as well as demonstrating that once the susceptible population falls below a certain threshold, either from natural or artificial immunization, the outbreak can no longer sustain itself and the disease will eventually burn itself out naturally. This would later lead to the concept of the basic reproduction number ($R0$) defined as the number of additional infections the average infected individual causes. When $R0 > 1$, the disease can perpetuate in the population, while if the $R0 \leq 1$, the disease will die out.

Since this initial study, compartmental models have remained a popular tool for modeling the dynamics of epidemics. While the underlying principle is the same, the model has evolved with time, such as dividing the population into heterogeneous risk groups who may be more or less likely to catch, transmit, or be negatively impacted by the disease [9, 44], or modeling the population as households [8, 31]. Another common alteration is the addition of new compartments, with some of the most common ones adding an "Exposed" group (SEIR) [25] which adds a latency period between catching a disease and being able to transmit it, and/or dividing "Recovered" into "Death" and "Recovered" (SEIRD) [5, 21, 47]. In recent years, there have also been several studies that add separate compartments for vaccinated and unvaccinated individuals (SVEIR) [44, 24, 9]. There are also SIS or SIRS models, where infected individuals eventually re-enter the pool of susceptible individuals to represent immunity falling over time.

A study by Gillis et al. [25] developed a simulation-optimization framework to help policymakers select the optimal response strategy to an outbreak of epidemic disease in order to minimize the total number of infected for a limited budget. The framework was based on an SEIR model, with additional compartments that divided the infected into many distinct categories based on the severity of an individual's symptoms (e.g., Mild-Moderate, Severe, Hospitalized, ICU) as well additional separate compartments for if an individual is isolating after being exposed to the disease

or not. The model simulated what effect different response strategies (closures, protection, travel restrictions) would have on the course of an outbreak and what cost it would incur on the economy of the region to maintain. They applied this framework to a case study in the Canadian province of Nova Scotia during the COVID-19 pandemic. They found that if the budget allows it, strict policies can work to severely limit the total cases, while if the budget is limited, oscillating between strict and relaxed closures is the best policy, and whenever possible to enforce mask-wearing and social distancing.

Recently, a new compartmental epidemiological model, called DELPHI, was developed by Li et al. [35], which was tailored specifically for predicting the cases and deaths for COVID-19 and the impact of different control measures. DELPHI is based on the mechanistic compartmental structure of SEIR models, with three major alterations. First, it adds the additional compartments of undetected, detected hospitalized, and detected quarantined. Second, measures of non-pharmaceutical interventions (NPIs) (social distancing, school closures, gathering restrictions, non-essential business restrictions, lockdowns) are factored into the disease activity (transmission and death rate). Finally, the model accounts for declining mortality rates over time, as medical practices improve. With this model, they found they were able to predict cases of and deaths from COVID-19 to within 6% and 11% accuracy, respectively, over a period of two weeks across 200 geographical areas. However, two shortcomings of DELPHI are: 1) the base DELPHI model does not account for vaccination, and 2) within regions, populations are homogeneous, and not divided into different risk classes. This can make it difficult to use DELPHI for vaccine allocation, as often specific sub-populations have different dynamics and need to be prioritized over one another. However, an extension on DELPHI, referred to as DELPHI-V, accounts for vaccination and heterogeneity across risk groups [9].

While compartmental models are far more common, in recent years with improvements in computational ability, agent-based simulation has begun to be used more frequently in epidemiological modeling. In agent-based simulation models, each individual is assigned attributes and adheres to a set of rules that impact the actions and interactions between individuals. Through these actions, interactions, and emergent behavior, the system is simulated. Dalgiç et al [15] compared the performance

of strategies derived from compartmental SEIR models and FluTe agent-based simulation across different scenarios and performed sensitivity analysis on both. The agent-based simulation was found to generally lead to lowered costs and cases compared to compartmental models, at the expense of being far more computationally expensive. While both models agreed on the importance of vaccinating school-age children, FluTe models were found to vary their strategy with slight differences in scenarios, while SEIR was less sensitive to change.

While the agent-based simulation approach has its benefits, its use in vaccine allocation studies continues to be rare in comparison to compartmental models. This is likely due to the simplicity and flexibility compartmental models bring, which allows them to be quickly adapted to new situations and scenarios, while not being overly sensitive to initial conditions and minor parameter changes. One problem with the compartmental model approach, or really any epidemiological modeling method, is that the number of newly infected is fundamentally a bilinear function of the current susceptible and infectious populations. This makes it easy to simulate the effect of various intervention strategies, but difficult to optimize their implementation. This model structure is what makes resource allocation problems so difficult when trying to control the dynamics of an epidemic disease.

2.3 Resource Allocation in Epidemic Scenarios

As previously mentioned, when examining existing work on the topic of how to allocate limited resources or implement control strategies to combat the spread of disease, the most common approaches are using analytical methods, simulation, and optimization techniques.

The nature of the problem seems to resist analytical methods. Its non-convex nature makes it difficult to derive useful insights beyond a few basic ones. However, Duijzer et al. [20] were able to use analytic methods to study SIR models and the impact of vaccine allocation on the spread of disease to derive new structural results and insights. First, through the analysis of the herd effect, they prove that the health benefits for a population as a function of the vaccination fraction are in general convex-concave and increasing-decreasing. Second, they investigate the relationship between administering a single vaccine and administering multiple ones and define

the concept of the dose-optimal vaccination fraction, the fraction that maximizes the health benefit per dose of vaccine in a population. At the dose-optimal vaccination fraction, the health benefits per dose decrease when moving away from this fraction in either direction. They argue that while the critical vaccination coverage, i.e., the coverage required to reach $R0 \leq 1$, is suitable for situations where there is a large supply of available vaccines, the dose-optimal fraction is the better guide for allocating scarce vaccines. However, the study does not consider interaction effects between different sub-populations, neither inter- nor intra-regional. Furthermore, it focused only on the best practice for a single initial allocation of vaccines and not the best practice for the allocation of vaccines over multiple periods, as otherwise finding the dose-optimal vaccination fraction becomes exceedingly difficult. The most useful takeaways from this work are the proof that the epidemiological benefits of vaccines are non-linear and that this effect can be measured.

While most compartmental models either assume a homogeneous population or divide the population into groups of individuals based on geographic regions and/or risk groups, a second approach is to partition the population into households. In household models, there are two levels of mixing: global and local. On the global level, each infectious individual has a small chance of passing the infection on to any susceptible individual, while on the local level, each infectious individual has a much higher probability of passing on the infection to other members of their local group or "household". Though frequently referred to as households, these local groups can actually be expanded to include neighborhoods or larger communities. This approach was first developed by Ball, Mollison, and Scalia-Tomba [8]. Through the analysis of the household model, they found that when this local level mixing was accounted for, it had an amplifying effect on the impact of an outbreak and the basic reproductive number, with a larger fraction of the population needing to be vaccinated to get an outbreak under control. From these results, they proposed that the optimal strategy with respect to minimizing the impact and duration of an outbreak is the equalizing strategy. The equalizing strategy allocates vaccines such that each household receives the same level of coverage. Expanding on this work, Ball and Lyne [7] showed that, at least under certain conditions, the equalizing strategy was also the optimal approach in regards to minimizing the overall cost of a vaccination program for a household

model. Further building off this work, Kneeling and Ross [31] use a household SIR model that accounts for differences in intra-household transmission rates and considers larger households to re-evaluate the performance of the equalizing strategy in different scenarios and identify where it breaks down. They found that both large household sizes and small within-household transmission rates would break the optimality of the equalizing strategy [7]. When this optimality breaks, they found there was no simple rule to determine the optimal distribution strategy. Another problem with household models is they can only provide insight into whether it is better to focus on fully vaccinating individual households or spread vaccines across the population. They have no means of distinguishing where or who should receive priority for vaccination. For these reasons, household models are not very common in other literature on the topic of vaccine allocation.

Choi and Shim [14] take a game theory approach to analyze the problem of vaccination. Using an extension on the SIR model, they developed a model that uses game theory to predict whether individuals will choose to partake in social distancing, vaccination, or both. Game theory works by expecting individuals to behave in a manner that is most advantageous to the individual, in this case, avoiding infection while incurring the least "cost" to themselves. From this, they can try to predict how individuals will act under different scenarios, such as what fraction of the population is likely to engage in social distancing or get vaccinated. This game theory approach is useful for predicting public uptake of vaccines, but not for designing a vaccination strategy to control an epidemic.

The common thread between these studies is that analytically deriving useful rules and relationships for the problem of controlling the spread of the epidemic disease has proven to be exceedingly difficult. The problem has to be simplified or abstracted to the point that it no longer resembles the original problem or system.

When studying the allocation of vaccines, the most common methodology is simulation or a combination of simulation and optimization techniques. Pure simulation models are built to quickly evaluate the performance of different strategies or evaluate a strategy under different model parameters and scenarios, rather than finding the optimal solution. For instance, Araz, Galvani, and Meyers [5] developed a variation on the traditional SEIR simulation model to include ineffectively and effectively

vaccinated groups, which they used to evaluate the effectiveness of different vaccination strategies under different scenarios. In this study, effectively vaccinated individuals moved to a "protected" group after an incubation period, during which they could still catch the disease, but after which they were considered completely immune and could no longer catch or transmit the disease. Meanwhile, ineffectively vaccinated individuals were unaware that they were not protected, thus did not seek re-vaccination, and were always susceptible to becoming infected. Using the model, they evaluated the effectiveness of Pro-rata, sequential by population, sequential-by-peak, and reverse sequential-by-peak vaccination strategies in minimizing cases and average wait time to be vaccinated. They found that the two best strategies were Pro-rata and reverse sequential-by-peak, with the latter having slightly fewer cases but longer average wait times versus Pro-rata, and both strategies made significant improvements from the base case of no intervention.

A similar study by Shim [63], applied a SEIR model to simulate the effects of different age-stratified vaccination strategies in South Korea during the COVID-19 pandemic. By testing numerous different potential vaccination strategies under varying supply and vaccine efficacy scenarios, they evaluated the outcome of each according to three different metrics: total infections averted, total deaths averted, and total Years of Life Lost (YLL). When there were enough vaccines to vaccinate at least 50% of the population, the best-performing strategy in terms of preventing infections focused on vaccinating adults between the ages of 20-49 years of age. This is because, in South Korea, this age bracket had the highest transmission incidence. When seeking to prevent deaths, the best-performing strategy shifted to instead prioritize adults older than 50. Then when preventing YLL, there was a slight shift compared to the preventing deaths strategy, as the best-performing strategy focused on those between 40-69 years old. While the strategies for preventing deaths and YLL were mostly consistent regardless of the supply of vaccines, as the vaccine supply became further limited, the best-performing strategies for preventing infections allocated more and more of the limited supply of vaccines to the 10-19 age bracket. These results are consistent with our own findings for the Ontario COVID-19 case study, discussed in Chapter 6.

In epidemic situations, the resources required to get an outbreak under control

are generally exponentially increasing with time. Rachaniotis, Dasaklis, and Pappis [57] tried to address this problem by combining a deterministic SIR simulation model with a single resource scheduling problem for medical teams that applies the concept of deteriorating jobs to represent the increasing time and effort required to contain an outbreak. They applied and evaluated their model against the case of mass vaccination against influenza A(H1N1) in Greece in 2009. 24 strategies were enumerated, simulated, and evaluated analytically for their performance. The best performance of these strategies was found to be able to significantly improve on the performance of the random allocation of medical teams.

While the previous examples have utilized simulation to predict the dynamics of disease under different parameters (such as the disease's transmission incidence, average infectious period, and contact rate), the reverse can also be done. Mukandavire et al. [47] developed and fit a SEIR simulation model to the cases of COVID-19 experienced in South Africa during the early stages of the pandemic. From this model, they were able to estimate the $R0$ for this early strain of COVID-19 ($R0 = 2.95$), as well as other infectious metrics such as the incubation and infectious periods. They also derived insights into the impact of lockdown on the number of contacts a person had in a day, observing an 80.31% reduction in contacts per day during lockdown, demonstrating what role such a reduction would play in slowing the spread of the disease. These estimates were derived extremely quickly, with the paper being accepted for publication in June of 2020, just months after COVID-19 had begun to spread around the world, providing valuable insights into the behavior of early COVID-19 which were corroborated by future studies [63, 35].

While using simulation to evaluate different vaccination strategies is extremely fast, the results are restricted to the test scenarios and strategies researchers develop synthetically. The model itself makes no decisions and there is no expectation of optimality. However, by combining the evaluation speed of simulation with optimization techniques, one can circumvent the difficulties that arise with trying to apply optimization techniques alone to the non-convex epidemiological model and use them to identify improved strategies over those that are developed manually. The usual process is to iterate between simulating the course of the disease and generating a vaccination strategy using optimization techniques. Using this iterative

process, simulation-based optimization can find near-optimal strategies tailored to the model's parameters and scenarios in a more reasonable time frame compared to using optimization alone.

Savachkin and Uribe [62] proposed a simulation-optimization approach to minimizing measures of morbidity, mortality, and societal and economic costs by distributing and redistributing resources such as vaccines, anti-viral drugs, and administrative capacity, over the course of an outbreak. The proposed approach consists of three models. The first handled cross-regional simulation of the spread of the outbreak between areas, to decide when an outbreak would spread to a region, while a separate set of models simulated the progression of the disease within each region. A separate optimization model was solved periodically to reallocate resources amongst the regions, while strategically keeping reserves for potential future outbreaks available. By applying their model to a hypothetical outbreak of H5N1 in Florida, they found that this proactive approach was able to significantly improve on a traditional reactive myopic policy. Liu and Zhang [38] presented a similar study where they used a dynamic logistics model to allocate various medical resources amongst geographic regions during an outbreak of epidemic disease. They believed previous work focused too much on a static problem without considering how the situation could evolve over time. Their work followed a closed forecast, plan, execute, and adjust the loop. First, a SEIR model would be run to forecast the course of the epidemic and the demand for medical resources. The forecast would then be used in a mixed 0-1 integer programming problem to identify the optimal allocation of medical resources to meet demand through a three-tiered logistics network (manufacturers, DCs, and clinics) for the lowest overall cost. Finally, the model's solution would be implemented in the real world and new data would be collected to adjust the forecast model for the next decision cycle. In both these studies (i.e., [62, 38]), the allocation of medical resources did not directly impact the course of the outbreak, and the epidemiological models are simply a forecasting tool for the demand. They focused on alleviating the worst effects of the disease, rather than trying to control it.

Conversely, Matter and Potgieter [44] used a combination of simulation and optimization to model the course of an epidemic, allocate medical teams, and distribute

vaccines via a drone network in an effort to directly control the course of an outbreak. The epidemic dynamics were simulated using a SEIRVD model. An integer programming problem is solved on each simulated day to maximize the expected prevented exposures (EPE) from allocating teams across the network, formulated as a knapsack problem. The allocation of drone deliveries was solved in a similar manner, by calculating the EPE value of delivery of vaccines and solving the knapsack problem with the additional constraint of limited daily delivery time. These EPE maximizing strategies were compared to and combined with several other strategies for allocating teams and deliveries such as simply prioritizing higher population areas or equally distributing resources across the region, with each combination of team and delivery strategy being run. They found that while for an untargeted vaccination campaign the EPE team allocation strategy performed poorly in terms of preventing exposures, the EPE delivery strategy outperformed all other delivery strategies. Meanwhile, for targeted vaccination campaigns, both the EPE team and delivery strategies performed extremely well, outperforming all other strategies. This EPE allocation approach shares a lot in common with one of the better-performing solution methods tested in our own study, namely the GMB heuristic.

Recently, Bertsimas et al. [9] used an extension on the DELPHI model [35] to optimize the location of vaccination sites and the allocation of vaccines to sites in order to minimize both deaths and the distances between vaccination sites and population centers. They made two augmentations to the base DELPHI model. The first was partitioning the population into risk classes to capture the disparate impact of the disease on different groups of people. The second was the capability to model the impact of vaccination on the dynamics of the pandemic. The model assumed there was only a single type of vaccine and there was no interaction between the different regions, with each region having its own separate DELPHI model. They also added smoothing constraints on the allocation of vaccines to prevent large fluctuations in the number of vaccines allocated to a region day to day. To combat the non-convexity issue of allocating vaccines, the authors devised a custom coordinate descent algorithm. This algorithm would iterate between simulating the dynamics of the disease under a certain vaccination strategy and then generating a new strategy

using a linear approximation of the bilinear non-convex equations from the simulation. A similar method, the fixed allocation linear approximation, was experimented with for solving our model, though the results proved to be less promising for the allocation of vaccines amongst regions and subgroups when compared to other methods explored. This was likely due to the simplicity of our own linear approximation generated from the simulation solution. An alternate linearization method would likely lead to an improved solution.

Pure optimization approaches are rare in the literature, as the non-convex nature of the problem makes it computationally difficult to solve. When it is used, usually some form of approximation is required.

Ng et al. [49] proposed a multi-criterion vaccination planning problem for seasonal influenza, with the objective of minimizing the vaccination cost and reproduction number while maximizing societal benefit. This model differed from other models in that rather than directly allocating vaccines, it sought to decide what portion of a population to vaccinate using a specific strategy, such as mass, targeted, and random. They used the model to identify the conditions under which each strategy was preferable. Targeted vaccination was found to perform better in the early stages of an outbreak, especially when there were limited vaccines, while mass vaccination became more favorable later on in an outbreak. The priority groups to vaccinate were found to be the elderly, who are most at risk from the disease, and school-age children, who are the most likely to spread the disease.

Enayati and Özaltın [21] presented an optimization model for finding the minimum vaccination coverage required to control an emerging outbreak of influenza across multiple geographic regions and sub-groups. The base model was a SEIR model with the addition of dividing the infected into quarantined and non-quarantined individuals. The decision variables are the portion of the population to be vaccinated in each sub-group (single period). Alongside the model, they proposed a solution algorithm that works by iteratively solving for upper and lower bounds by approximating the bilinear terms using discretization and exact linearization. Even with this algorithm, in a case study with only six regions and five age groups, the model took over two hours to solve to optimality. Through the model, they explored the effect of vaccine efficacy, enforcing equity across age groups and population centers,

and altering the probability of an infected individual isolating or not.

Majrajt et al. [41] developed an optimization method for the vaccine allocation problem based on the Nelder-Mead method and were able to observe significant improvement over other vaccination strategies. Particle swarm optimization was also used with similar results. While the authors refer to their solution as the optimal, it should be noted that for non-convex problems; as is the case with their underlying model, Nelder-Mead optimization can only locate a local optimum, and does not guarantee global optimality. Additionally, this work does not directly consider the allocation of vaccines and instead focuses on what fraction of different sub-groups of the population to vaccinate (single period). An extension on this work [42] applies their model and solution method to exploring the differences between the allocation strategies to minimize different disease metrics for one versus two-dose vaccines and investigate how parameters like vaccine efficacy, transmission rate, and supply impact the derived strategies.

The aforementioned studies (i.e., [49, 21, 41, 42]) do not consider vaccine allocation over time, and instead treat it as a one-time decision of deciding what fraction of the population to vaccinate, essentially when setting the initial conditions for the SEIR model. Multi-period vaccine allocation problems are even more uncommon in the literature. Of the papers that did use multi-period allocation, Ren, Órdoñez, and Wu [59] was one of the most comprehensive works, presenting two SIR optimization models tailored to modeling a smallpox outbreak in a single city and a network of cities, respectively. The objective of the models was to minimize deaths by deciding how to allocate a limited pool of vaccines to control an outbreak of smallpox. In addition to the usual parameters and constraints, their model had the ability to choose whether to adopt a ring or mass vaccination approach in a period and how many resources to dedicate to either strategy. The behavior of the ring vaccination strategy would even change depending on if mass vaccination had at any point been utilized in the city in the past. In their model, the usually non-convex SIR transmission rate is replaced with a constant transmission rate, based on a second-order Taylor approximation, to make it easier to solve. The authors found that for short time frames, with 4 time periods of 15 days each, constant-rate approximations would deviate by only 2% on average from those of the non-convex formulations, with a maximum

deviation of 11%. However, this result is specific to smallpox with a relatively low transmission rate, and would differ for other diseases.

Buhat et al. [12] took an alternate approach to linearize the SEIR model by using the maximum recorded R0 in a region to determine the maximum outbreak size as a function of the susceptible population. The number of susceptible individuals after vaccination is a function of the population size, the active cases at the time of vaccination, and the number of administered vaccines multiplied by their efficacy. The model seeks to minimize deaths for a given budget. They applied their model to the allocation of vaccines in the Philippines during the COVID-19 pandemic and found that while richer regions could afford to vaccinate their entire population, poorer regions could get the most for their limited resources by vaccinating 60-70% of the population.

While most papers on the subject of vaccine allocation primarily focus on controlling the course of the disease, there are some that also try to enforce some measure of equity into the model, either in the constraints or objective function, in order to guide aspects of the model's decision-making. Balcik et al. [6] took an alternate approach and is entirely focused on maximizing equity for a wide range of different regions and groups of people, while still ensuring reasonable coverage is achieved. Unlike the other studies covered thus far, this paper highlighted the importance of considering whether a vaccination strategy is equitable or not and whether some groups will be left vulnerable when deciding whether to implement it or not.

2.3.1 Stochastic & Robust Approaches

The previous papers took a deterministic approach to the problem of resource allocation in epidemic scenarios. Fu, Sim, and Zhou [24] argued that this deterministic approach often used in SEIR models was not realistic to the reality of disease spread and instead proposed a robust formulation of an SVEIR to optimally allocate vaccines on a limited budget. They made slight alterations to their model to linearize the constraint and make it easier to solve by modifying the calculation of the new exposure and replacing it with a normal approximation.

In addition to robust formulations, there have been studies that developed a stochastic approach to the problem. Examples of such studies include Yarmand et

al. [68], who proposed a two-phase stochastic simulation model for the allocation of seasonal influenza vaccines amongst disparate geographic regions. The first phase occurs before the beginning of flu season. A limited pool of vaccines was allocated amongst the different regions with the goal being to contain the spread of the disease for the lowest cost. The simulation then advanced to the second stage and determined whether the disease was successfully contained in each region as a stochastic function of the administered vaccination strategy. In regions where the disease was not contained, a second round of vaccines would be administered at an increased cost compared to the first phase. As such, the model's purpose was to carefully balance the risk of over-vaccinating against the risk of under-vaccinating. Structurally, this closely resembles the news-vendor model and the first phase decision can be solved as such.

Similarly, Nguyen and Carlson [50] developed a discrete stochastic simulation model to simulate the spread of infection and the real-time allocation of vaccines, aiming to complement and compare results with studies that used deterministic methods. They focused on the reactive allocation of vaccines and how the time delay to allocation could impact the course of the disease in a scenario where vaccines were allocated all at once in a single time period. They identified the trade-off between the number of available vaccines and the time to intervention on the spread of disease, outbreak duration, and final outbreak size, and demonstrated that with increasing delay in intervention, the quantity of vaccine required to keep an outbreak under control grows exponentially. They also investigated the impact of coupling cities together, demonstrating that the optimal vaccination strategy when considering coupled cities depends on the strength of the coupling. The best strategy identified was to exclusively focus on one city in the event of weak coupling and to move towards equal distribution as coupling grows stronger.

Yin and Büyüktahtakın [69] presented a stochastic multi-stage model to minimize total infected and required funerals across multiple regions through the distribution of both medical centers and resources under a limited budget. The model was built for use in outbreaks of Ebola, which primarily impacts low-income countries that need to make the most of limited resources. The model accounted for yet-to-be-buried bodies as Ebola requires careful and costly burials for the deceased in order to prevent the

risk of further infection. Their model is different from others discussed thus far as they linearize the transmission of disease in their model by modeling the course of the disease as the probability of a series of discrete scenarios occurring, allowing them to efficiently solve the model. They also include two forms of equity metrics: infection equity, enforcing a limit on the absolute deviation between regions relative to the number of infected, and capacity equity, enforcing a limit on the absolute deviation between the amount of treatment capacity established in a region relative to all other regions. The strategies developed by their model could significantly improve on existing strategies for combating Ebola outbreaks. By enforcing or relaxing these equity constraints on the model, they demonstrated that strict adherence to these equity constraints can be very costly in terms of additional infections and deaths.

Mohammadi et al. [46] presented both a stochastic and robust version of a model to optimize the design of a vaccine distribution network with the objective of minimizing total deaths and the cost of distribution and administration. Their network followed a four-tier hierarchical hub network design of national, regional, and departmental warehouses, with vaccines being administered at vaccination centers. Each vaccination center had a set of "demand points" that it was able to cover. The model tracked the cost of opening warehouses and vacation centers, placing an order, purchasing vaccines, transportation and holdings costs, and the cost of administering the vaccine to an individual. Each vaccination center was modeled as a multi-server GI/G/c queue that limited the number of individuals who could be vaccinated in a given period. Furthermore, these demand points were subdivided into different risk classes, with the risk class determining an individual's infection, hospitalization, and mortality risk, as well as what vaccines they were potentially eligible for and their number of contacts in a week. Different types of vaccines had different efficacy and costs, with some vaccines requiring up to two doses and a time delay between doses. The stochastic nature of the model comes from the chances of inventory disruptions, inventory degradation, capacity disruption, and capacity degradation over time, along with different scenarios for contact, infection, mortality rates, and vaccine effectiveness. The epidemiological side of this model was simplified to a linear approximation with each risk group having a static probability of becoming infected in any given period. The authors expressed an interest in trying to integrate their

model with a SIR model to account for the variable transmission rates over time that would be experienced in a real-world outbreak.

2.4 Vaccine Allocation and Distribution

It was found that studies that examine both the allocation and distribution of vaccines for the purposes of controlling and minimizing epidemics are relatively rare. As previously mentioned, Rastegar et al. [58] and Tavana et al. [66] decided how to best allocate and distribute a limited supply of vaccines when demand outstrips supply. However, they did not use any epidemiological modeling and instead had a static demand for vaccines. Meanwhile, Liu and Zhang [38] used a SEIR model to periodically forecast demand for medical supplies for which they then determined the optimal allocation to minimize distribution costs. While the SEIR model was being used to determine where to allocate resources, the allocation of those resources was reactive and had no impact on the course of the epidemic. Going a step further, Fadaki et al. [22] calculated the risk of being left unvaccinated for each client. They then used a transshipment model to determine how best to allocate and distribute vaccines to minimize this risk when there were not enough vaccines to cover demand. While a wide variety of factors went into calculating this risk, there was still no direct use of an epidemiological model, and they were not trying to impact the course of the disease. However, since the COVID-19 pandemic, there has been an increase in interest and studies on the topic, and there have been several studies that considered both the optimal allocation and distribution strategy to control an epidemic.

One such study was the previously mentioned Matter and Potgieter [44], which approached the problem as a knapsack problem. Using a SEIRVD simulation model, they solved a 0-1 integer programming model each day to allocate medical teams and vaccines where their allocation saw the most immediate benefit in terms of expected prevented exposures (EPE) for the following day. A similar approach was used to design the drone distribution network to deliver the vaccines to the medical teams. It was essentially using a GMB heuristic, allocating resources where individually the most immediate impact is seen, which cannot guarantee optimality. As such, when performance was compared to other strategies, the EPE strategy under-performed in certain scenarios.

Manupati et al. [40] used decision tree analysis to develop different vaccination strategies and evaluated them using the synthetic control method for the expected reduction in the number of cases and deaths from implementing each strategy. They then used a stochastic MILP to build a custom distribution network to fulfill the selected strategy. The model determined the location and number of cold storage facilities, their required inventory levels, and the allocation of vaccines to clinics, so as to minimize both transportation time and cost. However, while both vaccine allocation and distribution are considered, they are solved separately, rather than as a single problem.

Using their extension on the DELPHI model, Bertsimas et al. [9] optimized the location of mass vaccination sites and the allocation of vaccines amongst said sites and sub-populations so as to minimize deaths, hospitalizations, and quarantined individuals. To circumvent the non-convex nature of the model, they developed an algorithm to iterate between simulating the dynamics of the pandemic under a certain vaccination strategy and then generating a new strategy using a linear approximation of the bilinear non-convex equations from the simulation. However, the authors were primarily focused on the high-level strategic decision of where to locate vaccination sites, rather than on how to allocate vaccines. Additionally, they did not model the distribution network required to supply the chosen mass-vaccination sites with their required vaccines, and instead simply assumed that no matter how many vaccines are allocated to a location they could be delivered and administered.

Finally, Mohammadi et al. [46] developed a detailed model for the design of multi-vaccine distribution networks built to serve an ongoing epidemic situation. The model decided what warehouses and vaccine centers to open, when to place orders, what inventory levels to keep, and how to distribute the vaccines through the network. Additionally, the model was stochastic, with the chance for disruptions to the network to occur. However, they used a simplified linear approximation of the epidemiological model where each risk class had a flat probability of becoming infected in any given period. The authors expressed a desire to replace this linear approximation with a SIR model to improve the accuracy to which they can model the course and impact of an epidemic.

2.5 Literature Review Classification Table

Table 2.1 classifies the literature. A study's focus can be classified based on whether it considers distribution (Dist); how resources get to their destination, site location (Loc); deciding where to locate facilities (such as DCs, clinics, vaccination sites), epidemiological modeling (Epid); whether disease dynamics are modeled, and allocation (Alloc); deciding how resources should be divided amongst a population. Studies can be further classified by their solution method (whether the paper uses analytical (Analy), simulation blue(Sim), heuristic (Heur), or optimization (Opt) techniques), and what objective metric the paper focuses on, (such as minimizing economic (Econ) or environmental (Env) costs, maximizing societal benefits (Soc), or minimizing epidemiological impact (Epid)). Each paper can also be classified based on whether it is deterministic or accounts for uncertainty (Uncrt) in one form or another. The last column of the table contains the topic of the paper's case study, if it had one.

Table 2.1: Literature Review Classification Table

	Study Focus				Solution Method				Objective					
Study	Dist	Loc	Epid	Alloc	Analy	Sim	Heur	Opt	Econ	Env	Soc	Epid	Uncrt	Case Study
[5] Araz, Galvani, & Meyers 2012			✓	✓		✓						✓		Epidemic Influenza - Arizona
[6] Balcik et al. 2022	✓			✓				✓			✓			COVID-19 - Turkey
[7] Ball & Lyne 2002			✓	✓	✓			✓				✓		-
[8] Ball, Mollison, & Scalia-Tomba 1997			✓	✓	✓			✓				✓		-
[9] Bertsimas et al. 2021	✓	✓	✓	✓		✓	✓	✓				✓		COVID-19 - USA
[12] Buhat et al. 2021			✓	✓				✓			✓	✓		COVID-19 - Philippines
[13] Chen et al. 2014	✓	✓						✓	✓					EPI Vaccination - Niger
[14] Choi & Shim 2020			✓	✓	✓						✓	✓	✓	-
[15] Dalgic et al. 2017			✓			✓	✓					✓		-
[18] De Carvalho, Ribeiro, & Barbosa-Póvoa 2019	✓	✓						✓	✓	✓	✓			European Supply Chain
[19] Duijzer, Van Jaarsveld, & Dekker 2018			✓	✓	✓							✓		-
[21] Enayati and Özaltın 2020			✓	✓				✓			✓	✓		-
[22] Fadaki et al. 2022	✓			✓				✓				✓		COVID-19 - Australia
[24] Fu, Sim, & Zhou 2021			✓	✓				✓	✓			✓	✓	COVID-19 - New York
[25] Gillis et al. 2021			✓			✓		✓	✓			✓		COVID-19 - Nova Scotia
[28] Hovav and Herbon 2017	✓	✓						✓	✓		✓			Seasonal Influenza - Israel
[29] Hovav and Tsadikovich 2015	✓	✓						✓	✓		✓			Seasonal Influenza - Israel
[31] Keeling and Ross 2015			✓	✓	✓			✓				✓		-
[33] Kermack and McKendrick 1927			✓		✓							✓		-
[35] Li et al. 2022			✓			✓		✓				✓		-
[38] Liu and Zhang 2016	✓		✓	✓		✓		✓	✓		✓		✓	-
[36] Lim, Norman, and Rajgopal 2019	✓						✓	✓	✓					EPI Vaccination - Africa
[40] Manupati et al. 2021	✓	✓	✓	✓		✓		✓	✓			✓	✓	COVID-19 - India
[41] Matrajt et al. 2021 A			✓	✓			✓	✓				✓		COVID-19 - Washington
[42] Matrajt et al. 2022 B			✓	✓			✓	✓				✓		COVID-19 - Washington
[44] Matter and Potgieter 2021	✓	✓	✓	✓		✓	✓					✓		Measles - Niger
[46] Mohammadi et al. 2022	✓	✓	✓	✓				✓	✓			✓	✓	COVID-19 - France
[47] Mukandavire el al. 2020			✓			✓						✓		COVID-19 - South Africa
[49] Ng et al. 2018			✓	✓				✓				✓		-
[50] Nguyen and Carlson 2016			✓	✓		✓		✓				✓	✓	-
[56] Pishavee, Razmi, & Torabi 2014	✓	✓						✓	✓	✓	✓		✓	Vaccine Supply Chain - Iran
[57] Rechaniotis, Dasaklis, & Pappis 2012			✓	✓		✓						✓		H1N1 - Greece
[58] Rastegar et al. 2021	✓	✓		✓				✓	✓		✓			Influenza
[59] Ren, Ordóñez, & Wu 2013			✓	✓	✓			✓				✓		Smallpox Bio-terror Attack
[61] Saif and Elhedhli 2016	✓	✓				✓		✓	✓	✓				Ontario Cold Chain
[62] Savachkin and Uribe 2012	✓		✓	✓		✓		✓	✓		✓		✓	H5N1 - Florida
[63] Shim 2021			✓	✓		✓						✓		COVID-19 - South Korea
[64] Sripada et al. 2023	✓							✓	✓				✓	COVID-19 - India
[65] Sun, Andoh, & Yu 2021	✓					✓			✓	✓				COVID-19 - Norway
[66] Tavana et al. 2021	✓	✓		✓				✓	✓		✓			Influenza
[68] Yarmand et al. 2014				✓				✓	✓			✓	✓	Seasonal Influenza - North Carolina
[69] Yin and Büyüktahtakın 2021	✓	✓	✓	✓				✓	✓		✓	✓	✓	Ebola - Africa
Our work	✓		✓	✓		✓	✓	✓	✓			✓		COVID-19 - Ontario

3whitelightgrey

2.6 Literature Gaps

To the best of our knowledge, there is currently no known method that can solve the problem of vaccine allocation to minimize measures of epidemiological impact (infected, deaths, YLL, outbreak duration) to the guaranteed optimal in a reasonable time frame (i.e., a few days to a week at most, as the problem likely does not need to be solved regularly) in the case of a limited supply of vaccines. Given the non-convex nature of the problem, it is at least NP-hard [43, 55], meaning that depending on the size of the problem, it can take days, weeks, or even months to solve, even using state-of-the-art commercial solvers like GUROBI and BARON. As such, the existing works on the problem of vaccine allocation either linearize the model or take a heuristic solution approach, often combined with simulation. We implemented several novel solution methodologies, carefully evaluated them in terms of their solution quality and their execution time, and were able to identify three promising approaches to solve the problem to near optimality and devise an exact solution method that is suitable for small-size problems.

Additionally, to the best of our knowledge, few studies have utilized epidemiological modeling to both optimally allocate and distribute vaccines in an epidemic situation. Most vaccine supply chain and distribution studies use models with a fixed demand for vaccines or try to simply distribute as many vaccines as possible for a given budget. A few studies have used epidemiological modeling as a forecasting tool for the future demand for vaccines, but the allocation of the vaccines did not actually impact the course of the outbreak. Meanwhile, studies focused on vaccine allocation for the purposes of controlling epidemic disease either largely ignored costs and supply chain considerations, instead only limiting supply, or reduced the costs to just the cost of acquisition and administration. On top of that, many allocation studies only assumed a single-period allocation of vaccines, rather than a multi-period allocation which allocates vaccines over time, and/or are homogeneous

(single population or multiple disconnected populations), rather than heterogeneous (multiple interconnected groups).

Our work presents a novel combination of a SEIR epidemiological model and a network model for the purpose of vaccine allocation and distribution that enables balancing the effectiveness of a vaccination strategy against the logistical limitation of the assumed allocation strategy. The model is heterogeneous and capable of handling multi-period allocation for a limited supply of multiple types of single-dose vaccines. Alongside this model, we explore several novel solution methods, including four promising ones. This model and solution methods are applied to a case study in the Canadian province of Ontario during the COVID-19 pandemic in 2021 and used to derive useful insights into the best practices for vaccine allocation when looking to prevent cases and/or deaths.

Chapter 3

Problem Description and Mathematical Formulation

Consider a scenario where there has been an epidemic outbreak of a disease in a region. The population has little to no natural immunity to the disease, and if no intervention is taken, the disease will rapidly spread through the population. One effective measure to control the outbreak is mass vaccination. However, it is infeasible to vaccinate the entire population at once due to supply and/or administrative capacity constraints. Decisions need to be made on how to best make use of these limited resources.

The population at large can be divided into a set of population zones, indexed by $i \in I$, and stratified based on factors such as age or vulnerability to disease (such as high-vulnerability or immunocompromised individuals), indexed by $g \in G$. A Susceptible-Exposed-Infected-Removed (SEIR) compartmental model can be used to model and predict the dynamics of the epidemic. Consider that there is a limited supply of vaccines that becomes available every time period (e.g., week) $t \in T$, of which there are different vaccine types $v \in V$ with varying efficacy, cost, and logistical requirements. These vaccines need to be stored in and shipped from DCs $j \in J$ to the population zones. The aim is to determine the number of doses of each type of vaccine to administer to each group in every population zone in every time period, so as to minimize some impact of the epidemic, such as total exposures, while balancing against the costs required to enact such a strategy and respecting the constraints of the distribution system. These constraints might include a limited vaccine supply, minimum shipment sizes, capacity constraints for the DCs and population zones, access and equity considerations, etc.

3.1 Notation

Table 3.1 presents the notation used in the models.

Table 3.1: Notation

Sets:	
I	Set of population zones (e.g., Public Health Unit in Ontario Canada) indexed by i
J	Set of vaccine distribution centers (DC) indexed by j
G	Set of risk classes (e.g. age groups) indexed by g
V	Set of vaccine types indexed by v
T	Set of time periods indexed by t
Parameters:	
d_{jv}	The unit cost (per lot) for sourcing and shipping vaccine type v to DC j
c_{ijv}	The unit cost (per lot) for shipping from DC j to population zone i for vaccine type v
p_j	Maximum capacity that can be shipped from DC j in a single period, measured in terms of volume (cubic inches)
q_{vt}	Total quantity of vaccine v available in time period t
s_j	Maximum storage capacity of DC j, measured in terms of volume (cubic inches)
a_i	Maximum number of doses that can be administered in zone i per time period (administrative capacity).
η_v	Efficacy of vaccine v
b_v	Number of shots per a lot of vaccines v
κ_v	Storage requirement of one lot of vaccine v in cubic inches
$C_{gg'}$	Matrix for inter-group mixing in a single period of time
τ	Disease transmissibility coefficient (i.e., probability of infection given contact)
α_1	Average exposure duration (i.e. the diseases incubation period)
α_2	Average infection duration (i.e. how long an individual remains infectious)
β	Conversion factor between epidemiological impact and total shipping costs
S^0_{ig}	Initial susceptible population in zone i in risk group g
E^0_{ig}	Initial exposed population in zone i in risk group g
I^0_{ig}	Initial infected population in zone i in risk group g
R^0_{ig}	Initial removed population in zone i in risk group g
r^0_{jv}	Initial stockpile of vaccine v at DC j
w_{ig}	Weight of new exposures for each zone i and risk group g (i.e. Mortality Rate)
Decision variables:	
x_{ijvt}	Lots of vaccine v shipped from DC j to zone i in period t
y_{jvt}	Lots of vaccine v supplied to DC j in time period t
r_{jvt}	Lots of vaccine v in DC j at the end of period t
z_{igvt}	Number of people in group g in zone i that receive vaccine v in period t
S_{igt}	Number of susceptible cases of group g in zone i at the end of period t
E_{igt}	Number of exposed cases of group g in zone i at the end of period t
I_{igt}	Number of infected cases of group g in zone i at the end of period t
R_{igt}	Number of removed cases of group g in zone i at the end of period t
NE_{igt}	New exposures in group g in zone i in period t

3.2 SEIR Epidemiological Model

The SEIR epidemiological model handles the disease dynamics and how an outbreak spreads through the population over time. The population of the region, already

broken down by population zone i and risk group g, is divided further into four compartments, denoted by S_{igt}, E_{igt}, I_{igt} and R_{igt}. These compartments denote, respectively, the number of susceptible, exposed, infected, and removed individuals in zone i belonging to risk group g at the end of time period t. While some SEIR models are continuous time, using differential equations, our model is instead of a discrete time and uses difference equations. Assuming a closed system and that the time period under consideration is not sufficient that natural population growth and migration will have a major impact, the total population $P_{ig} = S_{igt} + E_{igt} + I_{igt} + R_{igt},\ (\forall i \in I, g \in G, t \in T)$, remains constant.

Within each zone, inter-group mixing happens according to a symmetric $|G| \times |G|$ contact matrix C, where $C_{gg'}$ specifies the average number of close contacts a person in group g has with people from group g' in a single period of time. Given a disease transmissibility coefficient (i.e., probability of infection given contact) τ, the number of new exposures, NE_{igt}, in period $t+1$ due to contact during period t is calculated as

$$NE_{ig(t+1)} = \tau S_{igt} \sum_{g' \in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}}, \tag{3.1}$$

where the term inside the summation is simply the product of the number of contacts by the probability that the contacted person is infectious, returning the number of contacts with infected people in group g'. Summing over all groups, it is then multiplied by the probability of infection given contact and the susceptible population of group g to find the number of new exposures in the group ($NE_{ig(t+1)}$). This number is then subtracted from the susceptible compartment (S_{igt}) and added to the exposed compartment (E_{igt}) to get their corresponding values in period $t+1$. Additionally, in every period, the flows from E to I and from I to R, are determined by the α_1 and α_2 parameters, which denote the average exposure and infection duration respectively, in the same units as t (e.g., weeks). The initial conditions, how many susceptible,

exposed, infected, and removed there are for each risk group g in zone i at the start of $t = 1$, are set using the parameters S^0_{ig}, E^0_{ig}, I^0_{ig}, and R^0_{ig}.

The decision variable z_{igvt} denotes the number of doses of vaccine of type v administered in zone i to risk group g in each time period t. A set Z can be defined by all the feasibility constraints imposed on the allocation of vaccines, such as supply chain and administrative limitations. We assumed that only the unvaccinated can contract and spread the disease, and thus the effectively vaccinated portion of each group ($\sum_{v \in V} \eta_v z_{igvt}$) is subtracted from S_{igt} and added to R_{igvt}, where η_v is the efficacy of vaccine v. Ineffectively vaccinated individuals simply remain in the susceptible compartment. If instead, we wanted to have vaccinated individuals still contract and spread the disease, just at a different rate, additional compartments would need to be added to represent the vaccinated populations (i.e., $VS_{igt}, VNE_{igt}, VE_{igt}, VI_{igt}$).

The objective function should seek to find $z \in Z$ in such a way that minimizes some epidemiological metric (total cases, hospitalizations, deaths, YLL). For this model, it was decided to focus on minimizing the total number of new exposures NE_{igt} (i.e., cases), and thus the size of the outbreak. However, a weight parameter w_{ig} can be used to adjust model behavior. For example, setting the weight to the mortality rate of each group g will have the model minimize deaths rather than cases. Alternatively, it can be used with separate vulnerable groups, such as the immunocompromised and frontline workers, to prioritize those groups for vaccination. Putting this all together, the SEIR model can be formulated as follows:

$$\min_{z\in Z,S,E,I,R,NE} \quad \sum_{i\in I}\sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} \tag{3.2}$$

$$\text{s.t.} \quad NE_{ig(t+1)} = \tau S_{igt} \sum_{g'\in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.3}$$

$$S_{ig(t+1)} \geq S_{igt} - NE_{ig(t+1)} - \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.4}$$

$$E_{ig(t+1)} = \left(1 - \frac{1}{\alpha_1}\right) E_{igt} + NE_{ig(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.5}$$

$$I_{ig(t+1)} = \left(1 - \frac{1}{\alpha_2}\right) I_{igt} + \frac{E_{igt}}{\alpha_1} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.6}$$

$$R_{ig(t+1)} = R_{igt} + \frac{I_{igt}}{\alpha_2} + \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.7}$$

$$S_{ig0} = S^0_{ig} \qquad \forall i \in I, g \in G \tag{3.8}$$

$$E_{ig0} = E^0_{ig} \qquad \forall i \in I, g \in G \tag{3.9}$$

$$I_{ig0} = I^0_{ig} \qquad \forall i \in I, g \in G \tag{3.10}$$

$$R_{ig0} = R^0_{ig} \qquad \forall i \in I, g \in G \tag{3.11}$$

$$S_{igt}, E_{igt}, I_{igt}, R_{igt}, NE_{igt} \geq 0 \qquad \forall i \in I, g \in G, t \in T \tag{3.12}$$

Constraint (3.3), based on (3.1), determines the number of new exposures for the next time period. Constraints (3.4) through (3.7) determine the dynamics of the population, i.e., how many are in the susceptible, exposed, infected, and removed compartments for each zone, group, and time period. Constraints (3.8) through (3.11) determine the initial conditions of the model. Constraint (3.12) enforces non-negativity.

There are numerous potential alterations that could be made to improve the realism of the model. Such as incorporating inter-zone mixing of the population, which might be modeled using constraints similar to the equation $\sum_{i'\in I} \delta_{ii'gt} = \sum_{i'\in I} \delta_{i'igt}$, where $\delta_{ii'gt}$ denotes the population of group g moving from zone i to i' in time period t. However, inter-zone mixing would account for a very small portion of the overall population and contacts, especially in an epidemic scenario where travel is likely to be kept to a minimum, and is thus over-shadowed by intra-zone mixing as soon

as the disease becomes established in a zone. Furthermore, as the model currently assumes that, once vaccinated, a person is unable to either contract or spread the disease, additional compartments could be built to represent the vaccinated population, separate from the rest of the population ($S'_{igt}, E'_{igt}, I'_{igt}$), where the vaccinated might have lower contract, transmission, and mortality rates from the disease, but are still a contributing factor to the epidemic. For diseases such as COVID-19, it might be useful to explore a model where every person requires multiple doses of a vaccine in order to be fully vaccinated. This could be achieved in a similar manner to incorporating vaccination into the model; by adding separate compartments to capture the first-dose and second-dose populations. This would also require dividing the vaccine allocation decision variables z_{igvt} into first dose allocation and second dose allocations. However, as even the single-dose vaccine model would prove difficult to solve, in this thesis the focus is on single-dose vaccines.

3.3 Vaccine Supply Chain Model

The vaccine supply chain model can be formulated as a standard network problem. Vaccines are either manufactured locally or imported from abroad. Let d_{jv} be the cost associated with sourcing and shipping one lot of vaccine type v to DC j, and q_{vt} be the total quantity available (in lots) of vaccine v in time period t. It can be assumed that, within a reasonable time frame, the quantity of available vaccines is determined ahead of time, based on production and shipping plans from the vaccine supplier(s), and thus are beyond the control of the decision-maker. We assume that there is always going to be some limiting factor on the vaccine supply in any time period. The decision variable y_{jvt} denotes the number of lots of vaccine v shipped to DC j in each time period t, and the vaccine lot size (doses per lot) is set by b_v. Each DC j can ship only a limited quantity of vaccine during each time period, denoted as p_j. It is assumed that the DCs are equipped to store vaccine lots for prolonged

periods (i.e., longer than the planning horizon) and turnover is high enough that perishability is not a major concern. The maximum storage capacity in DC j is s_j and each vaccine has a storage space requirement κ_v, both of which are measured in cubic inches. The actual inventory of vaccine v in DC j at the end of period t is denoted as r_{jvt}, with the initial inventory levels being r^0_{jv}. The unit (lot) shipping cost from DC j to population zone i of vaccine type v is c_{ijv}. Finally, x_{ijvt} is the lots of vaccine v shipped to population center i from DC j in each time period t. With that, for a given demand for vaccines $\sum_{g\in G} \overline{z}_{igvt}$ the vaccine supply chain problem can be formulated as:

$$\min_{x,y,r} \quad \sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv}y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv}x_{ijvt} \tag{3.13}$$

$$\text{s.t.} \quad \sum_{j\in J} y_{jvt} \le q_{vt} \qquad \forall v \in V, t \in T \tag{3.14}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \le p_j \qquad \forall j \in J, t \in T \tag{3.15}$$

$$b_v \sum_{j\in J} x_{ijvt} \ge \sum_{g\in G} \overline{z}_{igvt} \qquad \forall i \in I, v \in V, t \in T \tag{3.16}$$

$$r_{jv(t+1)} - r_{jvt} = y_{jv(t+1)} - \sum_{i\in I} x_{ijv(t+1)} \qquad \forall j \in J, v \in V, t \in T \tag{3.17}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \le s_j \qquad \forall j \in J, t \in T \tag{3.18}$$

$$r_{jv0} = r^0_{jv} \qquad \forall j \in J, v \in V \tag{3.19}$$

$$y_{jvt}, r_{jvt}, x_{ijvt} \ge 0 \qquad \forall i \in I, j \in J, v \in V, t \in T \tag{3.20}$$

$$y_{jvt}, r_{jvt}, x_{ijvt} \in \mathbb{Z} \tag{3.21}$$

The objective function (3.2) seeks to minimize the shipping and sourcing costs, both from the suppliers to the DCs and from the DCs to the population zones. Constraint (3.14) states that only as many vaccines as are available can be sourced and shipped to the DCs in each time period, while (3.15) states that the volume of vaccines distributed from each DC cannot exceed a set volume. Constraint (3.16)

requires that the vaccines shipped from all DCs must collectively meet the demand for vaccines in a region. Constraint (3.17) is the flow balance constraint for the DCs, which states that the change in inventory from one period to the next equals the difference between received and shipped vaccine quantities. Constraint (3.18) limits how much can be stored in a DC while (3.19) defines the initial inventory levels of the DCs. Constraint (3.20) defines the domain of the decision variables.

There are numerous potential additions or alterations that could be made to the supply chain model to improve on its accuracy. For instance, a common structure for vaccine supply chain models is a multi-tier network, with vaccines starting from a large stockpile and working their way down through a series of DCs. As it currently is formulated, the model is essentially a three-echelon model (Supply, DC, Zone), but additional intermediate stages could be added relatively easily without major modifications. Alternatively, rather than stopping at zones, the model could go a layer deeper and show how vaccines are distributed to individual clinics and vaccination sites. Taking it a step further, an alternate approach entirely from the simple transshipment model could be to treat the distribution of vaccines as a vehicle routing problem. However, for this thesis, we elected to keep the vaccine supply chain model simple yet effective. It serves the goal of limiting the supply of vaccines and loosely modeling how they might be distributed. It is not meant to accurately predict the true vaccine distribution costs, but instead, be quick to solve while providing a rough estimate of the sourcing and shipping costs. The complexity of this problem arises almost entirely in solving the vaccine allocation portion of the problem, with the distribution side being far easier comparatively. In fact, so long as the supply chain model eventually connects to the SEIR model through a constraint such as (3.16), where demand represents the chosen allocation of vaccines, all solution methods explored in this paper can be applied with only minor alterations.

3.4 Integrating the Models

The epidemiological and supply chain models can be combined into a single multi-objective model where the epidemiological portion tries to minimize the epidemiological impact through the optimal allocation of vaccines and the supply chain portion imposes limits on vaccine availability and seeks to minimize the monetary costs of a given vaccine allocation strategy. These two objectives are adversarial as administering more vaccines reduces epidemiological impact but increases costs, and vice versa. Using a conversion factor of β, the model can be made to prioritize minimizing cases or costs. β essentially denotes how much the decision-maker is willing to spend in order to prevent one additional unit of the chosen epidemiological impact metric (new exposures, deaths, YLL, etc.). The combined model is as follows:

$$\min_{S,E,I,R,NE,z,x,y,r} \quad \sum_{i\in I}\sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} \tag{3.22}$$
$$+ \beta(\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv} y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv} x_{ijvt})$$

$$\text{s.t} \quad NE_{ig(t+1)} = \tau S_{igt} \sum_{g'\in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.23}$$

$$S_{ig(t+1)} \geq S_{igt} - NE_{ig(t+1)} - \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.24}$$

$$E_{ig(t+1)} = \left(1 - \frac{1}{\alpha_1}\right) E_{igt} + NE_{ig(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.25}$$

$$I_{ig(t+1)} = \left(1 - \frac{1}{\alpha_2}\right) I_{igt} + \frac{E_{igt}}{\alpha_1} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.26}$$

$$R_{ig(t+1)} = R_{igt} + \frac{I_{igt}}{\alpha_2} + \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{3.27}$$

$$\sum_{j\in J} y_{jvt} \leq q_{vt} \qquad \forall v \in V, t \in T \tag{3.28}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \leq p_j \qquad \forall j \in J, t \in T \tag{3.29}$$

$$r_{jvt} - r_{jv(t-1)} = y_{jvt} - \sum_{i\in I} x_{ijvt} \qquad \forall j \in J, v \in V, t \in T \tag{3.30}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \leq s_j \qquad \forall j \in J, t \in T \tag{3.31}$$

$$\sum_{g\in G} z_{igvt} \leq b_v \sum_{j\in J} x_{ijvt} \qquad \forall i \in I, v \in V, t \in T \tag{3.32}$$

$$\sum_{g\in G}\sum_{v\in V} z_{igvt} \leq a_i \qquad \forall i \in I, t \in T \tag{3.33}$$

$$S_{ig0} = S0_{ig} \qquad \forall i \in I, g \in G \tag{3.34}$$

$$E_{ig0} = E0_{ig} \qquad \forall i \in I, g \in G \tag{3.35}$$

$$I_{ig0} = I0_{ig} \qquad \forall i \in I, g \in G \tag{3.36}$$

$$R_{ig0} = R0_{ig} \qquad \forall i \in I, g \in G \tag{3.37}$$

$$r_{jv0} = r0_{jv} \qquad \forall j \in J, v \in V \tag{3.38}$$

$$z_{igvt}, S_{igt}, E_{igt}, I_{igt}, R_{ijt}, NE_{igt}, x_{ijvt}, y_{jvt}, r_{jvt} \geq 0 \tag{3.39}$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \in \mathbb{Z} \tag{3.40}$$

Constraints (3.23)-(3.27) are the SEIR constraints of the epidemiological model, which determine the disease dynamics. Constraints (3.28)-(3.31) are the supply chain constraints. Constraint (3.32) is the only constraint that connects the two models together, by enforcing that the collective vaccines delivered to a zone must be greater than or equal to the administered vaccines. Constraint (3.33) is a new constraint that enforces that the number of vaccines administered in a zone must be less than or equal to the administrative capacity of the zone. For the sake of simplicity, we assume administrative capacity is constant over time but depends on the geographical zones where the vaccines are being administered. However, a_i could be replaced with another equation, such as increasing administrative capacity over time to represent increased efficiency and practice in administering vaccines, or this constraint could also be replaced with a similar constraint

$$\sum_{j \in J} \sum_{v \in V} b_v x_{ijvt} \leq a_i \qquad \forall i \in I, t \in T, \tag{3.41}$$

which connects the amount of vaccines delivered to a region to the administrative capacity, rather than the amount that is used. This would prevent wasting partial lots of vaccines (e.g., sending 2 lots of 1000 vaccines but only having the capacity to use 1300 of them, wasting 700 vaccines). This constraint would be useful if decision-makers wanted to avoid wasting any vaccines. This problem could also be corrected by allowing the zones to store a small number of vaccines on-site between time periods, thus allowing them to store a partial lot of vaccines for use in the next period if they exceed their administrative capacity for the current period. As is, the model will avoid wasting vaccines, unless the benefits outweigh the incurred costs. Constraints (3.34)-(3.38) set the initial conditions of the model. Constraint (3.39) is the non-negativity constraint, and constraint (3.40) states that x_{ijvt},y_{jvt}, and r_{jvt} must be integers.

As previously mentioned, this model is a multi-objective one, balancing epidemic

impact against monetary costs using a conversion factor β. Alternatively, we could add an additional constraint to the model that imposes a cap on the total cost (i.e., a budget). This would remove the requirement to select a specific value for β and would allow decision-makers to define precisely how much money is available for the vaccination campaign. This would also circumvent the requirement to attach a monetary cost to prevented deaths, which can be a sensitive matter, both morally and politically. The Pareto frontier between the epidemiological impact and costs of both the budget-constrained and multi-objective models would be identical, as by setting the budget to the optimal cost incurred under a specific value of β, the optimal allocation strategy would be identical.

Chapter 4

Solution Methods

On the surface, this model may seem relatively easy to solve. For the SEIR epidemiological portion of the model, the majority of the model is just a series of difference equations simulating the dynamics of the epidemic based on the vaccine allocation strategy represented by the decision variable z_{igvt}. The remaining decision variables (S_{igt}, E_{igt}, I_{igt}, R_{igt}, NE_{igt}) are state variables determined explicitly from this value. For the supply chain portion, while a mixed integer problem, it is relatively simple to solve as a transshipment model, where values of the x_{ijvt}, y_{jvt}, and r_{jvt} decision variables again rely on the value of the decision variable z_{igvt}, as it represents the demand for vaccines in the model. However, the real complexity arises with constraint (3.23), which can be shown to not only be non-linear, but non-convex in nature. This bilinear constraint and the inclusion of the bilinear term NE_{igt} in constraints (3.24) and (3.25) makes the solving process significantly more complex.

Because where vaccines are administered directly impacts the susceptible population (3.24), which in turn impacts the new exposures (3.23), as well as the future susceptible and infected populations, where every vaccine will be most effective depends on where every other vaccine is administered. As such, deciding where to administer a vaccine is a nearly endlessly recursive decision. Consider also that the secondary objective of minimizing costs means that not every vaccine will even necessarily be administered and that there are potentially multiple different strategies that can arrive at the same objective value just with a different balance between costs and new exposures. Hence, it becomes apparent why the vaccine allocation and distribution problem has proven so difficult to solve to optimality.

4.1 Commercial Solvers

It was found that commercially available solvers, such as GUROBI v10 or BARON v22, were ineffective for all but the simplest of problems. In general, if the number of zones is greater than five, the number of groups is greater than three, or if there are more than ten time periods or one type of vaccine, a commercial solver will be unable to solve the model to optimality or even near optimality in a reasonable time frame (i.e., takes longer than 24 hours.). The optimality gap is often very large, relatively speaking, often stalling for hours at a gap of $\geq 50\%$. Even when the solver is warm-started using the best solution obtained from another solution method, it is usually unable to further improve on the provided solution. This is not unexpected, as most commercial solvers are not built to efficiently solve non-convex problems. However, for the small problems that commercial solvers can solve, they can provide a point of comparison for other solution methods.

4.2 Lagrangian Relaxation

The first solution method that was explored was using Lagrangian relaxation [23]. Lagrangian relaxation works by relaxing the "complicating" constraints to decompose the problem into sub-problems. Without the complicating constraints, the sub-problems are, in theory, easier to solve, with the violation of the relaxed constraint being linearly penalized in the objective function according to the Lagrangian multiplier. By solving the sub-problems, a Lagrangian bound (lower bound) on the optimal solution can be obtained. Then, by using a method such as Kelley's cutting plane method [32], additional cuts can be generated from the sub-problem solutions for use in the dual master problem, which will provide both new Lagrangian multipliers and an upper bound on the optimal solution. These new Lagrangian multipliers are plugged back into the sub-problems and this process is repeated until convergence.

4.2.1 Relaxed Model

The model below is the relaxed combined supply chain and epidemiological model, where constraint (3.32) was relaxed.

$$\min_{S,E,I,R,NE,z,x,y,r} \quad \sum_{i\in I}\sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} \tag{4.1}$$
$$+\beta\left(\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv}y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv}x_{ijvt}\right)$$
$$+\sum_{i\in I}\sum_{v\in V}\sum_{t\in T}\lambda_{ivt}\Big(\sum_{g\in G} z_{igvt} - b_v\sum_{j\in J} x_{ijvt}\Big)$$
$$\text{s.t} \quad \text{Constraints} \quad (3.23)-(3.31)$$
$$\text{Constraints} \quad (3.33)-(3.40)$$

This method was initially attempted because by relaxing constraint (3.32), which is the only constraint joining the supply chain and the epidemiological parts of the model, using the Lagrangian multiplier λ_{ivt}, the problem can be decomposed into a supply chain sub-problem (SP1) and $|I|$ different epidemiological sub-problems (SP2), which can all be solved individually.

4.2.2 SP1

Based on the supply chain model, SP1 (4.2) is a MILP that is always bounded and returns a solution feasible to the original formulation. It is always bounded as even if the term $(\beta c_{ijv} - \lambda_{ivt} b_v)$ is negative, constraints (3.29)-(3.30) impose an upper bound on the value of x_{ijvt}. It can be solved quickly as a transshipment model using commercial solvers.

$$\theta_1 = \min_{x,y,r} \quad \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T}(\beta c_{ijv} - \lambda_{ivt} b_v)x_{ijvt} + \sum_{j\in J}\sum_{v\in V}\sum_{t\in T}\beta d_{jv} y_{jvt} \tag{4.2}$$

$$\text{s.t} \quad \text{Constraints} \qquad (3.28) - (3.31), (3.38)$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \geq 0 \;\&\; \in \mathbb{Z}$$

4.2.3 SP2

SP2 is based on the SEIR epidemiological model and can be further decomposed into $|I|$ identical sub-problems. Unfortunately, each sub-problem is still non-convex due to constraint (3.23), which still makes SP2 significantly difficult to solve using commercial solvers. Therefore, it is necessary to limit the number of risk groups, vaccine types, and time periods in order to solve SP2. Even with these smaller problems, it is sometimes only possible to obtain an upper bound on the sub-problem rather than the true optimal, which severely impacts performance. In theory, the sub-problem could be solved to optimality through simulation and enumeration if the problem is reduced to a single risk group (homogeneous model) rather than using multiple risk groups (heterogeneous model) and some limit is imposed on the number of vaccines that can be allocated to a zone. This limit could be something like the initial susceptible population or the overall supply of vaccines. Alternatively, by solving SP1 first, the results from it could be plugged into SP2 to limit the supply of vaccines, which ensures the solution to SP2 is always feasible for the original problem. SP2 can thus be formulated as (4.3)

$$\theta_{2i} = \min_{S,E,I,R,NE,z} \quad \sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} + \sum_{g\in G}\sum_{v\in V}\sum_{t\in T} \lambda_{ivt} z_{igvt} \tag{4.3}$$

$$\text{s.t} \quad \text{Constraints} \quad (3.23) - (3.27)$$

$$\text{Constraints} \quad (3.34) - (3.37)$$

$$z_{igvt}, S_{igt}, E_{igt}, I_{igt}, R_{igt}, NE_{igt} \geq 0 \qquad v \in V, g \in G, t \in T \tag{4.4}$$

4.2.4 Dual Master Problem

For the dual master problem, the objective function (4.5) seeks to maximize $\theta_1 + \sum_{i\in I} \theta_{2i}$. K is the index set of feasible solutions obtained by solving the sub-problems. Constraints (4.6) and (4.7) are cuts derived from previous iterations of solving the sub-problems. Constraint (4.8) is included as without it, for certain values of λ_{ivt}, the value of θ_{2i} can become unbounded. This occurs when by decreasing θ_1, the problem is able to increase the values of θ_{2i} such that it will always result in a net increase in the objective function. Therefore, (4.8) is included to prevent θ_{2i}, representing total new exposures in zone i, from surpassing the total population of this zone, which would be both infeasible to the original problem and illogical to consider as a potential solution. The dual master problem is a simple LP and thus can be solved efficiently using commercially available solvers.

$$\max_{\lambda_{ivt},\theta_1,\theta_{2i}} \quad \theta_1 + \sum_{i\in I} \theta_{2i} \tag{4.5}$$

$$\text{s.t} \quad \theta_1 \leq \Big(\sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} (\beta c_{ijv} - \lambda_{ivt} b_v) x^k_{ijvt} + \beta \sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv} y^k_{jvt}\Big) \qquad \forall k \in K \tag{4.6}$$

$$\theta_{2i} \leq \sum_{g\in G}\sum_{t\in T} w_{ig} NE^k_{igt} + \sum_{g\in G}\sum_{v\in V}\sum_{t\in T} \lambda_{ivt} z^k_{igvt} \qquad \forall i \in I, k \in K \tag{4.7}$$

$$\theta_{2i} \leq \sum_{g\in G} P_{ig} \qquad \forall i \in I \tag{4.8}$$

$$\lambda_{ivt} \geq 0 \tag{4.9}$$

4.3 Linear Approximation

Another approach to deal with the non-convexity in the epidemiological model is to replace the bilinear constraints with a linear approximation before solving. This might lead to a different solution from the one obtained from using the original formulation. However, if the linear approximation is a close approximation of the bilinear constraint or is used as part of a larger solving algorithm, such as alternating between optimizing a linear approximation and simulating the non-convex model, then potentially a near-optimal solution could be obtained from it.

4.3.1 McCormick Envelope

Lets define the variable $Q_{igg't} = S_{igt}I_{ig't}$. Constraint (3.23) can then be rewritten as follows:

$$NE_{ig(t+1)} = \tau \sum_{g' \in G} C_{gg'} \frac{Q_{igg't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T \tag{4.10}$$

Given that the original variables have some known lower and upper bound, i.e., $I^l \leq I \leq I^u$ and $S^l \leq S \leq S^u$, the bilinear term $Q_{igg't} = S_{igt}I_{igt}$ can be approximated using the following inequalities [60]:

$$\begin{aligned}
Q_{igg't} &\geq S^l_{igt}I_{ig't} + S_{igt}I^l_{ig't} - S^l_{igt}I^l_{ig't} && (4.11)\\
Q_{igg't} &\geq S^u_{igt}I_{ig't} + S_{igt}I^u_{ig't} - S^u_{igt}I^u_{ig't} && (4.12)\\
Q_{igg't} &\leq S^u_{igt}I_{ig't} + S_{igt}I^l_{ig't} - S^u_{igt}I^l_{ig't} && (4.13)\\
Q_{igg't} &\leq S^l_{igt}I_{ig't} + S_{igt}I^u_{ig't} - S^l_{igt}I^u_{ig't} && (4.14)
\end{aligned}$$

These equations collectively form the McCormick Envelope [60] and provide an approximation on the value of the bilinear term, which allows the model to be solved as a linear model. The quality of this approximation depends on the tightness of the bounds on I_{igt} and S_{igt}.

4.3.2 MC Heuristic

Observe that the complicating decision variables are S_{igt} and I_{igt}. The model can thus be linearized by fixing either of these variables and optimizing the other. Additionally, by fixing S_{igt} to $\overline{S}_{igt}$, the allocation of vaccines z_{igvt} can no longer impact the model, as its impact on the model is exclusively to lower the value of S_{igt+1}. Thus when fixing S_{igt}, z_{igvt} must also be fixed to $\overline{z}_{igvt}$ or the model will not allocate any vaccines as they do not impact the new exposures and only serve to increase costs. To simplify the model, the removed compartment R_{igt} can be dropped as the removed population has no impact on the outcome of the model. The fixed susceptible model can be formulated as follows:

$$\min_{E,I,NE,x,y,r} \quad \sum_{i\in I}\sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} + \beta\left(\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv} y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv} x_{ijvt}\right) \tag{4.15}$$

$$\text{s.t} \quad NE_{ig(t+1)} = \tau \overline{S}_{igt} \sum_{g'\in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.16}$$

$$E_{ig(t+1)} = \left(1 - \frac{1}{\alpha_1}\right) E_{igt} + NE_{ig(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.17}$$

$$I_{ig(t+1)} = \left(1 - \frac{1}{\alpha_2}\right) I_{igt} + \frac{E_{igt}}{\alpha_1} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.18}$$

$$\sum_{j\in J} y_{jvt} \leq q_{vt} \qquad \forall v \in V, t \in T \tag{4.19}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \leq p_j \qquad \forall j \in J, t \in T \tag{4.20}$$

$$r_{jvt} - r_{jv(t-1)} = y_{jvt} - \sum_{i\in I} x_{ijvt} \qquad \forall j \in J, v \in V, t \in T \tag{4.21}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \leq s_j \qquad \forall j \in J, t \in T \tag{4.22}$$

$$\sum_{g\in G} \overline{z}_{igvt} \leq b_v \sum_{j\in J} x_{ijvt} \qquad \forall i \in I, v \in V, t \in T \tag{4.23}$$

$$E_{ig0} = E^0_{ig} \qquad \forall i \in I, g \in G \tag{4.24}$$

$$I_{ig0} = I^0_{ig} \qquad \forall i \in I, g \in G \tag{4.25}$$

$$r_{jv0} = r^0_{jv} \qquad \forall j \in J, v \in V \tag{4.26}$$

$$E_{igt}, I_{igt}, NE_{igt}, x_{ijvt}, y_{jvt}, r_{jvt} \geq 0 \tag{4.27}$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \in \mathbb{Z} \tag{4.28}$$

By fixing I_{igt} to $\overline{I}_{igt}$, both the Exposed and Removed compartments no longer impact the model's behavior and can be dropped to further simplify the model. The fixed infected model can then be formulated as follows:

$$\min_{S,NE,z,x,y,r} \quad \sum_{i\in I}\sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} + \beta\left(\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv} y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv} x_{ijvt}\right) \tag{4.29}$$

$$\text{s.t} \quad NE_{ig(t+1)} = \tau S_{igt} \sum_{g'\in G} C_{gg'} \frac{\overline{I}_{ig't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.30}$$

$$S_{ig(t+1)} \geq S_{igt} - NE_{ig(t+1)} - \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.31}$$

$$\sum_{j\in J} y_{jvt} \leq q_{vt} \qquad \forall v \in V, t \in T \tag{4.32}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \leq p_j \qquad \forall j \in J, t \in T \tag{4.33}$$

$$r_{jvt} - r_{jv(t-1)} = y_{jvt} - \sum_{i\in I} x_{ijvt} \qquad \forall j \in J, v \in V, t \in T \tag{4.34}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \leq s_j \qquad \forall j \in J, t \in T \tag{4.35}$$

$$\sum_{g\in G} z_{igvt} \leq b_v \sum_{j\in J} x_{ijvt} \qquad \forall i \in I, v \in V, t \in T \tag{4.36}$$

$$\sum_{g\in G}\sum_{v\in V} z_{igvt} \leq a_i \qquad \forall i \in I, t \in T \tag{4.37}$$

$$S_{ig0} = S^0_{ig} \qquad \forall i \in I, g \in G \tag{4.38}$$

$$r_{jv0} = r^0_{jv} \qquad \forall j \in J, v \in V \tag{4.39}$$

$$z_{igvt}, S_{igt}, NE_{igt}, x_{ijvt}, y_{jvt}, r_{jvt} \geq 0 \tag{4.40}$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \in \mathbb{Z} \tag{4.41}$$

By alternating between solving the problem while fixing S_{igt} or I_{igt}, it was conjectured that these approximations would work together to gradually converge on a stable solution that is a good approximation of the optimal solution to the original non-convex problem.

4.3.3 Fixed Allocation Linear Approximation

Alternatively, by fixing z_{igvt} and using simulation to solve the SEIR portion of the model the values of NE_{igt} and S_{igt} for a given value of z_{igvt} can be obtained ($\overline{NE}_{igt}$, $\overline{S}_{igt}$). The bilinear constraint (3.32) can then be replaced with a linear constraint that calculates NE_{igt} as a function of the value of S_{igt}.

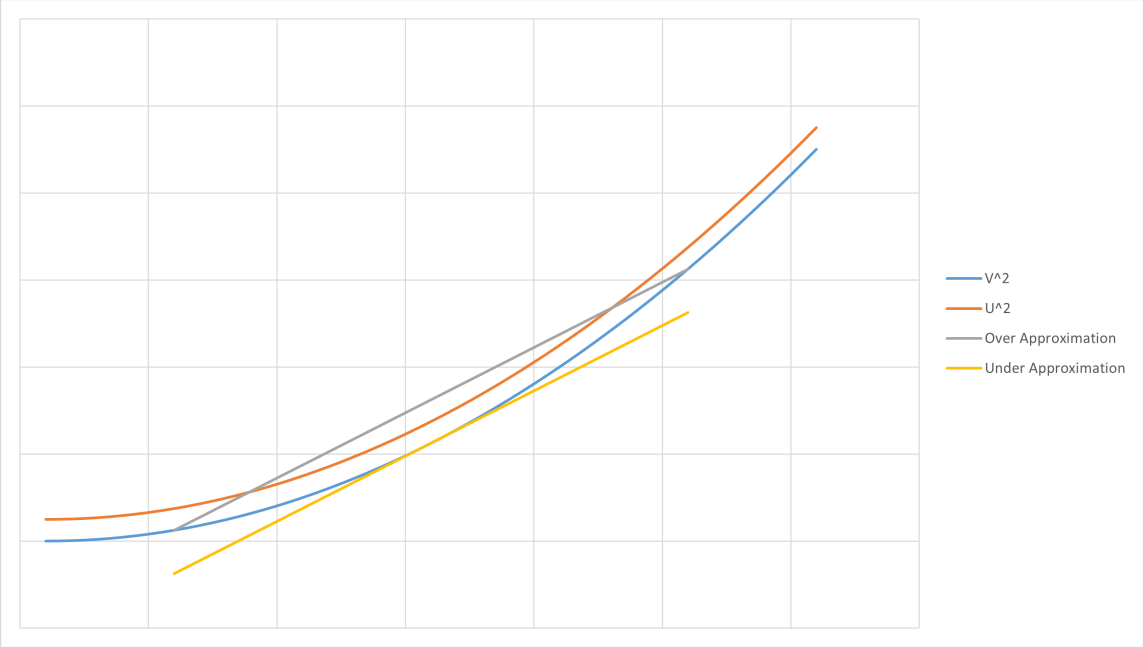
$$NE_{igt+1} = \frac{\overline{NE}_{igt+1}}{\overline{S}_{igt}} S_{igt} \qquad \forall i \in I, g \in G, t \in T \backslash \{t_n\} \tag{4.42}$$

Observe that for the same value of z_{igvt} this linear constraint is equivalent to constraint (3.32). By solving this linear model without fixing z_{igvt}, a new value for z_{igvt} can be obtained, which can be used to get new values for $\overline{NE}_{igt}$ and $\overline{S}_{igt}$ using simulation. Similar to the previous method, it was conjectured that these approximations would gradually converge to a stable solution that is a good approximation of the optimal solution to the original problem.

4.3.4 Piecewise-Linear Approximation

Rather than converting the non-convex constraint (3.32) into a single linear approximation, it can be approximated using a piecewise-linear function. Let $Q_{igt} = \tau \sum_{g' \in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}}$. Then, (3.32) can be written as $NE_{ig(t+1)} \geq S_{igt} Q_{igt}$. Next, define the variables $U_{igt} = (S_{igt} + Q_{igt})/2$ and $V_{igt} = (S_{igt} - Q_{igt})/2$, which enables the constraint to be written as $NE_{ig(t+1)} \geq U_{igt}^2 - V_{igt}^2$. Although the right-hand side is a non-convex difference-of-convex (d.c.) function, it can be easily piece-wise linearly approximated. In fact, only the second term (V_{igt}^2) needs to be approximated since

Figure 4.1: Piecewise-Linear Approximation Over and Under Approximation Visual Aid



it is the concave term.

There are two potential approaches to handling a piece-wise-linear approximation of V_{igt}^2, either an over-approximation or an under-approximation. An over-approximation would not work for this problem without a very large number of breakpoints. This is because when substituted back into $NE_{ig(t+1)} \geq U_{igt}^2 - V_{igt}^2$, an over-approximation can sometimes allow V_{igt}^2 to take on a value greater than U_{igt}^2, which would mean there would be no new exposures, as seen in figure 4.1.

Instead, an under-approximation must be used. Initially, it was speculated that the under-approximation could be generated iteratively over time as a series of linear cuts, however, this ended up not working as intended. Instead, we use a series of breakpoints. Because Q_{igt} is strictly positive, the maximum value $V_{igt} = (S_{igt} - Q_{igt})/2$ can take on is $S_{ig}^0/2$. Thus V_{igt}^2 can be approximated using

a series of k breakpoints ($\overline{v}^k_{igt}$) between 0 and $S^0_{ig}/2$, and a set of SOS2 decision variables (μ^k_{igt}). The more breakpoints used, the tighter the approximation to the original. By generating breakpoints $\sqrt{\epsilon}$ apart starting from 0 until surpassing $S^0_{ig}/2$, the maximum deviation from the original when drawing a straight line between any two points is limited to ϵ. Any value of V_{igt} can then be linearly approximated as $\sum_{k\in K}(\overline{v}^k_{igt})^2\mu^k_{igt}$. This piecewise-linear approximation of the SEIR model can then be divided by population zone I and expressed as follows:

$$\min_{S,E,I,R,NE,z} \quad \sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} \tag{4.43}$$

$$\text{s.t} \quad NE_{ig(t+1)} \geq U_{igt}^2 - \sum_{k\in K}((\bar{v}_{igt}^k)^2 - \epsilon)\mu_{igt}^k \qquad g\in G, t\in T\backslash\{t_n\} \tag{4.44}$$

$$Q_{igt} = \tau \sum_{g'\in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}} \qquad g\in G, t\in T \tag{4.45}$$

$$U_{igt} = (S_{igt} + Q_{igt})/2 \qquad g\in G, t\in T \tag{4.46}$$

$$\sum_{k\in K} v_{igt}^k \mu_{igt}^k = (S_{igt} - Q_{igt})/2 \qquad g\in G, t\in T \tag{4.47}$$

$$\sum_{k\in K} \mu_{igt}^k = 1 \qquad g\in G, t\in T \tag{4.48}$$

$$S_{ig(t+1)} \geq S_{igt} - NE_{ig(t+1)} - \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad g\in G, t\in T\backslash\{t_n\} \tag{4.49}$$

$$E_{ig(t+1)} = \left(1 - \frac{1}{\alpha_1}\right) E_{igt} + NE_{ig(t+1)} \qquad g\in G, t\in T\backslash\{t_n\} \tag{4.50}$$

$$I_{ig(t+1)} = \left(1 - \frac{1}{\alpha_2}\right) I_{igt} + \frac{E_{igt}}{\alpha_1} \qquad g\in G, t\in T\backslash\{t_n\} \tag{4.51}$$

$$R_{ig(t+1)} = R_{igt} + \frac{I_{igt}}{\alpha_2} + \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad g\in G, t\in T\backslash\{t_n\} \tag{4.52}$$

$$S_{ig0} = S_{ig}^0 \qquad g\in G \tag{4.53}$$

$$E_{ig0} = E_{ig}^0 \qquad g\in G \tag{4.54}$$

$$I_{ig0} = I_{ig}^0 \qquad g\in G \tag{4.55}$$

$$R_{ig0} = R_{ig}^0 \qquad g\in G \tag{4.56}$$

$$z_{igvt}, S_{igt}, E_{igt}, I_{igt}, R_{igt}, NE_{igt} \geq 0 \qquad v\in V, g\in G, t\in T \tag{4.57}$$

$$\mu_{igt}^k \geq 0, SOS2 \tag{4.58}$$

$$z\in Z \tag{4.59}$$

where μ_{igt}^k are special ordered sets of type 2 variables, of which at most 2 consecutive can take nonzero values and v_{igt}^k are the breakpoints selected in the range $[0, S_{ig}^0/2]$.

4.4 Greedy Marginal Benefit Heuristic

A promising approach to solving the problem of vaccine allocation is using a greedy heuristic to maximize the immediate marginal benefits. This method takes advantage of the fact that while finding the optimal vaccine allocation is difficult, for any given vaccination strategy (z_{igvt}), the outcome can be solved for efficiently using simulation. By starting from the baseline scenario of sending no vaccines, and then simulating the outcome of sending some portion of the vaccines to each zone and group combination, the marginal benefit for each potential allocation can be measured. By choosing to allocate a given portion of vaccines where the most immediate marginal improvement is observed, and repeating this process until all vaccines have been allocated, a very good solution can be found relatively quickly. This algorithm takes from a few seconds for simple problems to just over 24 hours for more complex scenarios.

4.4.1 Proportioning Scheme

The proportioning scheme used determines how the supply of vaccines is divided up into N portions. This is done for each type of vaccine v and for each time period t $(P_1^{vt}, \ldots, P_N^{vt})$, with each portion being allocated one at a time. There are countless potential methods that could be used to proportion the vaccine supply. In this work, three different proportioning schemes were experimented with.

The first is to allocate one vaccine at a time. This is the slowest possible method, requiring potentially millions or even billions of potential allocations to be evaluated to arrive at a solution, depending on the supply of vaccines and the number of zones and groups. The one benefit of this approach is it generally leads to slightly better solutions through a more granular search of the solution space, though this is not always the case, especially as the value of β increases.

The second approach is to allocate in bulk, such as portions of base ten units (e.g., 1000, 100, 10, 1). Starting from allocating 1000 doses of vaccine at a time,

when there are 10,000 doses left, start allocating 100 at a time. When there are 1000 left, allocate 10 at a time, and when there are 100 left allocate 1 at a time. This base ten scheme is not special, and the supply could just as easily be divided according to some other pattern, but base ten is convenient. A shortcoming of this approach though is there is the potential that if care is not taken in how the portions are created, not every solution is possible. Consider a scenario where there are 1,000 doses, divided into 9 portions of 100 doses, 9 portions of 10 doses, and 10 1 dose portions, with 15 possible zone and group combinations to potentially allocate to. Under the previously described proportioning scheme, it would be impossible to achieve a nearly equal distribution of the supply of vaccines, as 9 groups would be allocated the first 900 doses, and there would only be 100 doses left over to spread across the remaining 3 groups. Regardless of whether an equal distribution is the best or not, using this proportioning scheme cannot even be considered a possibility. This can be avoided by creating smaller portion sizes, however, if there are too many portions the run time will begin to increase again.

Therefore, a third approach is to divide the supply into as few portions as possible while still being able to achieve every possible solution through some combination of the portions, which can be done as follows. To start, assuming that the entire supply will be used, the largest portion of the supply of vaccines that can be guaranteed any zone and group combination will receive is equal to the amount received if the supply were to be divided equally amongst all zones and groups, rounded down to the nearest integer.

$$P_1^{vt} = \lfloor q_{vt}/(|I| * |G|) \rfloor. \tag{4.60}$$

This portion of the vaccines can then be removed from the supply pool, and the remaining supply divided equally again amongst all zones and groups.

$$P_K^{vt} = \lfloor (q_{vt} - \sum_{k=1}^{K-1} P_k^{vt})/(|I| * |G|) \rfloor. \tag{4.61}$$

This proportioning process repeats until $(q_{vt} - \sum_{k=1}^{K-1} P_k^{vt})/(|I| * |G|) \leq 1$, at which point, the remaining supply of vaccines is divided into portions of size 1. Through different combinations of these portions of the supply of vaccines, any possible allocation can be achieved.

While the second and third methods significantly improve the solve time of the algorithm; as there are fewer portions to be allocated, because they allocate in such large portions the granularity of the search is compromised and it can lead to worse solutions compared to allocating one at a time. This could be partially corrected by using more portions than strictly necessary while still limiting the number of portions that need to be allocated. However, in certain situations, increased granularity can also lead to worse solutions compared to a less granular search. This will be discussed in more detail in section 5.3.6.

4.4.2 Allocation and Measuring Marginal Benefit

Once the supply is divided into portions, a simulation is run to get the epidemiological impact for the case where no vaccines are administered; the distribution costs of such a solution would necessarily be zero and thus are ignored for now. This serves as the baseline solution from which the marginal benefit will be measured.

Starting from the first time period and the first type of vaccine, simulations are run of the SEIR epidemiological model for the case where $P_1^{1,1}$ is individually allocated to each zone and group combination to get the epidemiological impact. For each set of groups in a zone, the best epidemiological solution is identified. The allocation of vaccines across all zones can then be translated into a demand for vaccines in lots and plugged into the transshipment model, which is solved efficiently using a commercial solver, such as GUROBI. There is no need to solve the transshipment

model for each group, as the distribution costs are the same for sending vaccines to a region, regardless of how they are allocated once they arrive.

The epidemiological impact and distribution costs are then combined to evaluate the solution quality. This is compared to the best currently identified solution quality. If it is better, this solution becomes the new best solution. Once all zone and group combinations are tested for $P_1^{1,1}$, the allocation of $P_1^{1,1}$ is locked to the best-identified zone and group combination and the algorithm will begin the allocation of the next portion ($P_2^{1,1}$), building off the previous one. In this way, vaccines are allocated so as to maximize the marginal benefit of each portion. If none of the potential allocations for the tested portion improve on the current best solution, that portion is discarded. This occurs when the cost of distributing that portion of vaccines outweighs its benefits. By starting from the feasible solution of NI and sufficiently punishing infeasible allocations, the heuristic is guaranteed to return a feasible solution.

Once all portions for a given vaccine v are either allocated or discarded, it will advance to the next vaccine type $v+1$ ($P_N^{v,t} \rightarrow P_1^{v+1,t}$). Once all vaccines have been allocated for time period t, the algorithm will advance to the next time period $t+1$ ($P_N^{V,t} \rightarrow P_1^{v,t+1}$). The algorithm terminates once all allocations up to $|T|-1$ have been performed. There is no need to explore the final time period as allocating vaccines in the final period can never have a positive impact on performance. The best-identified allocation strategy is then extracted and its performance, decision variables' values, along with the algorithm run time are recorded.

4.4.3 Model Structure and Pseudocode

A pseudocode for simulating the epidemiological impact of a given vaccine allocation strategy can be found in Algorithm 1 and the pseudocode of the GMB heuristic can be found in Algorithm 2.

The supply chain model requires minor alterations in order to work with the GMB heuristic. Because the allocation is generated separately from the supply chain

Algorithm 1 SEIR Simulation Algorithm Pseudocode

Accepts z_{igvt}

$S_{ig0} \leftarrow S^0_{ig}$, $E_{ig0} \leftarrow E^0_{ig}$, $E_{ig0} \leftarrow E^0_{ig}$, $R_{ig0} \leftarrow R^0_{ig}$ ▷ Set Initial Conditions

for $t = 1 : T$ **do**

  for $i = 1 : I$ **do**

    for $g = 1 : G$ **do**

      for $g' = 1 : G$ **do** ▷ Calculate # of Infected Contacts

        $Contacts_{igt} = Contacts_{igt} + C_{gg'} I_{ig't} / P_{ig'}$

      end for

      $NE_{igt+1} = \tau S_{igt} Contacts_{igt}$ ▷ Calculate New Exposures

      $TotalNE = TotalNE + w_{ig} NE_{igt+1}$ ▷ Track Total New Exposures

      for $v = 1 : V$ **do** ▷ Calculate "Effective" Vaccines Administered

        $Vaccines_{igt} = Vaccines_{igt} + \eta_v z_{igvt+1}$

      end for

      $S_{igt+1} = max(0, S_{igt} - NE_{igt} - Vaccines_{igt})$ ▷ Calculate Susceptible

      $E_{igt+1} = max(0, (1 - (1/\alpha_1))E_{igt} + NE_{igt+1})$ ▷ Calculate Exposed

      $I_{igt+1} = max(0, (1 - (1/\alpha_2))I_{igt} + E_{igt}/\alpha_1)$ ▷ Calculate Infected

      $R_{igt+1} = R_{igt} + I_{igt}/\alpha_2 + Vaccines_{igt}$ ▷ Calculate Recovered

    end for

  end for

end for

Return $TotalNE, NE_{igt}, S_{igt}, E_{igt}, I_{igt}, R_{igt}$

constraints, it has the potential to be infeasible to the original problem, such as allocating more vaccines than there is the capacity to administer or distribute. To fix this, the constraints are relaxed, and instead, violations are penalized in the objective function. Table 4.1 presents the notation used in the relaxed supply chain model. The modified supply chain model used alongside the GMB heuristic is formulated as follows.

Table 4.1: Notation for Relaxed Supply Chain Model

Parameters:	
D_{ivt}	Demand in batches for vaccine v in region i at time period t
AP	Penalty incurred for each additional unit of administrative capacity required
CP	Penalty incurred for each additional unit of logistical capacity required
SP	Penalty incurred for each additional batch of vaccine required
Decision variables:	
ap_{it}	Units of additional administrative capacity required in region i during time period t
cp_{jt}	Units of additional logistical capacity required in DC j during time period t
sp_{vt}	Lots of additional vaccine v supply required to satisfy demand in time period t

$$\min_{x,y,ap,cp,sp} \quad \beta \left(\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv}y_{jvt} + \sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv}x_{ijvt} \right) \tag{4.62}$$
$$+ \sum_{i\in I}\sum_{t\in T} APap_{it} + \sum_{j\in J}\sum_{t\in T} CPcp_{jt}$$
$$+ \sum_{v\in V}\sum_{t\in T} SPsp_{vt}$$

$$\text{s.t} \quad \sum_{j\in J} y_{jvt} \leq q_{vt} + sp_{vt} \qquad \forall v \in V, t \in T \tag{4.63}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \leq p_j + cp_{jt} \qquad \forall j \in J, t \in T \tag{4.64}$$

$$r_{jvt} - r_{jv(t-1)} = y_{jvt} - \sum_{i\in I} x_{ijvt} \qquad \forall j \in J, v \in V, t \in T \tag{4.65}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \leq s_j \qquad \forall j \in J, t \in T \tag{4.66}$$

$$\sum_{j\in J} x_{ijvt} \geq D_{ivt} \qquad \forall i \in I, v \in V, t \in T \tag{4.67}$$

$$\sum_{v\in V} b_v D_{ivt} \leq a_i + ap_{it} \qquad \forall i \in I, t \in T \tag{4.68}$$

$$r_{jv0} = r^0_{jv} \qquad \forall j \in J, v \in V \tag{4.69}$$

$$x_{ijvt}, y_{jvt}, r_{jvt}, ap_{it}, cp_{jt}, sp_{vt} \geq 0 \tag{4.70}$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \in \mathbb{Z} \tag{4.71}$$

4.5 Evolutionary Genetic Algorithm

The heuristic algorithm described in the previous section demonstrates that simulating the outcome and performance of any given vaccination strategy z_{igvt} can be done very efficiently. What the heuristic lacks, being a greedy algorithm, is the ability to thoroughly explore the solution space. Due to the non-convex nature of the problem, this greedy search tends to lead to a local rather than the global optimum. In order to combat this, an Evolutionary GA [16] approach was used. Starting from a pool of

Algorithm 2 GMB Pseudocode

$z_{igvt} \leftarrow 0$
$BestNE = inf$
for $t = 2 : T - 1$, $v = 1 : V$ **do**
    $Q = 0$
    $portion = 10^{\lfloor log_{10}(b_v q_{vt}) \rfloor - 1}$
    while $Q < b_v * q_{vt}$ **do**
        if $(b_v * q_{vt}) - Q < 10 * portion$ **then**
            $portion = max(1, int(portion/10))$
        end if
        for $i = 1 : I$ **do**
            for $g = 1 : G$ **do**
                $z_{igvt} = z_{igvt} + portion$
                $Epi.Impact \leftarrow$ Run SEIR Simulation(z_{igvt})
                if $Epi.Impact < BestEpi.Impact$ **then**
                    $BestEpi.Impact = Epi.Impact$
                    $BestZg = z$
                end if
                $z_{igvt} = z_{igvt} - portion$
            end for
            Convert $BestZg$ (Doses) into $Demand$ (Lots)
            $Costs \leftarrow$ Solve Supply Chain Model for $Demand$
            $SolutionPerformance = BestEpi.Impact + Costs$
            if $SolutionPerformance < BestSolution$ **then**
                $BestSolution = SolutionPerformance$
                $BestZ \leftarrow BestZg$
            end if
        end for
        $Q = Q + portion$
        $z \leftarrow BestZ$
    end while
end for
Return BestZ, Best Solution

random vaccination strategies, through crossover and random mutation, better solutions are continuously generated, allowing for a wider search of the solution space than the GMB heuristic.

4.5.1 Chromosomes

For this GA, each chromosome is a four-dimensional array of genes of size $|I| \times |G| \times |V| \times |T|$. Each gene in a chromosome represents how many vaccines of type v are allocated to any given group g in zone i in each time period t. This is equivalent to the previously defined z_{igvt} decision variable that linked the two parts of the model and represents a potential vaccination strategy that could be utilized to control the epidemic.

4.5.2 Fitness Function

The fitness function of the GA takes in a chromosome (z) and evaluates performance in terms of the expected epidemiological impact and distribution costs associated with that chromosome's vaccine allocation strategy. The fitness function uses the same SEIR simulation (Algorithm 1) and relaxed supply chain model (equations (4.62)-(4.70)) used for evaluating the quality of vaccination strategies in the GMB heuristic. The relaxed supply chain model is still required as any given chromosome might represent an infeasible solution to the original problem. The fitness function first solves for epidemiological impact and then converts the individual vaccines in a zone into demand in lots, which is used in solving the transshipment model. The combined epidemiological impact and distribution costs are the chromosome's fitness score.

4.5.3 Crossover

Crossover is the process of taking two chromosomes, referred to as the "parent" chromosomes, and creating a new chromosome from them, referred to as the "child" chromosome. The purpose of crossover is to introduce new combinations of existing genes into the population. The crossover process works as follows:

First, the parent chromosomes are selected. Generally speaking, the parent chromosomes should be selected from the top-performing chromosomes of a given generation, rather than the bottom. This will lead to better-performing genes perpetuating and poorer-performing genes dying out. However, simply taking the top two performing chromosomes in each generation would limit the pool of available genes and limit the exploration of the solution space. Instead, it is better to give preference to the selection for higher performers, but still allow worse-performing chromosomes to occasionally be selected. The weights for each chromosome are generated using equation (4.72), where N is the total number of chromosomes and n is a chromosome's position in a list sorted by its fitness score.

$$weight_n = (N - n)^2 \tag{4.72}$$

A random number is generated between 0 and the sum of the weights. By looping through and subtracting the weight of each chromosome from the random number in order until the random number is smaller than the currently selected chromosome weight, higher-weighted chromosomes will be selected more often than lower ones.

Only the top X% of chromosomes have a chance of being selected as a parent, as determined by the parent selection range parameter, i.e., if set to a value of 0.3, the GA will only consider the top 30% of chromosomes as potential candidates for parent selection, preferring to select the highest performing chromosomes amongst them. This is done as the number of good chromosomes is vastly outweighed by bad chromosomes, especially as the solution begins to approach the limits imposed on it

by the supply chain constraints, and many of the worse-performing chromosomes are infeasible solutions.

Additionally, whenever a parent chromosome is selected, there is a random chance that it will be replaced with a predetermined "extreme" chromosome. An extreme chromosome would be a solution such as sending nothing or sending the maximum administrative capacity of a zone, for every gene in a chromosome. This is done because even if these chromosomes perform poorly in terms of their own fitness, they can be used to introduce these extreme genes back into the population, where they might serve to improve the solution and help maintain genetic diversity. Specifically, the extreme chromosome of sending nothing has proven to be useful in expediting the run time, as in many scenarios, the optimal strategy sees many genes take on a value of zero. By introducing the extreme chromosome, the GA is able to efficiently mix these genes into the population where they will be most effective, rather than relying solely on a mutation to introduce genes with a value of zero in the right location.

Once the two parent chromosomes are determined, genes are picked at random and their values are swapped, creating two new children in the process. Genes are only swapped with genes that occupy the same position in a chromosome, and there is no swapping of genes between groups, regions, vaccine types, or time periods. To what degree the two solutions are mixed, i.e., what ratio of their genes are swapped, is controlled by the mix ratio parameter. A mix ratio of 0.5 would mean each child contains half the genes of each parent, while a mix ratio of 0.2 (or 0.8) means each child has 1/5 of the genes of one parent and 4/5 of the other. Through experimentation on test data, it was found that a mix ratio of 0.3 was most effective for our problem.

4.5.4 Mutation

"Mutation" is the process of randomly introducing variation into the population in order to seed potential improvement. In each generation, after crossover is performed, there is a random chance that a chromosome will be chosen for mutation, creating a copy of the chosen chromosome and altering a gene at random.

Due to the large number of genes present in a chromosome ($|I| \times |G| \times |V| \times |T|$), even mutating 5% of the genes in a chromosome would result in mutations on numerous genes. Too many mutations make it less likely that helpful mutations will be able to stand out from unhelpful mutations. For example, gene X is mutated in such a way that it would improve the fitness score of a chromosome. However, at the same time gene Y is also mutated in such a way that it would harm the fitness score by the same or more than the previous mutation. This helpful mutation will likely be lost as its chromosome's fitness score is penalized overall. It was found that limiting the mutation rate to just one gene per chromosome improved the performance of the algorithm, especially as the algorithm nears the optimal solution, and mutations become increasingly likely to generate infeasible solutions.

There are two different methods of mutation in this GA. The first method is altering a randomly selected gene by a random amount between -X and +X, where X is the mutation size parameter. If such a mutation would change a gene to have a value less than or equal to 0, it instead takes on a value of 0. This form of mutation is useful for varying the total number of vaccines being allocated. In theory, just this form of mutation and crossover could find the optimal solution given sufficient time and generations. However, due to the fact that the fitness function penalizes solutions that violate feasibility or send more vaccines than required when solutions are near the optimal, it becomes increasingly unlikely for mutation or crossover to generate a better feasible solution on their own. Sending more vaccines results in a penalty, sending fewer vaccines results in a loss of benefit. Therefore, the only way to improve would be, through mutation and crossover, to randomly create a

chromosome where the exact same amount of vaccines were taken from one group, while that same amount of vaccines were given to another. This is very unlikely. Even more unlikely that this alteration will be beneficial. To combat this, a special kind of alternate mutation function, called "recombination", was introduced.

"Recombination" is an alternate form of mutation that accounts for how the specific problem of vaccine allocation is structured. Recombination works by causing mutations where if one group would have vaccines taken away, a second group would receive those vaccines, and vice versa. This allows the algorithm to explore solutions where the same amount of vaccines were being used, just in different ways. Recombination works by selecting a "From" gene and a "To" gene at random in a chromosome. It only makes sense to re-allocate vaccines in this manner if the selected genes are of the same vaccine type and time period, so the re-combination function first chooses the "From" gene and then based on the vaccine type and time period of that gene picks the "To" gene. A random percentage of the value of the "From" gene is then taken and transferred to the "To" gene. Additionally, when a gene has a value of 0, it is impossible to send anything from it, and so if a gene with a value of 0 is chosen as the "From" gene the recombination would be wasted. To prevent this, the algorithm will check whether the randomly selected "From" gene has a value of 0 and if it does, it will continue to pick random genes until it picks a gene that does not have a value of 0. This is especially helpful because the optimal solution is likely to have numerous genes with a value of 0, which would lead to many useless mutations if not accounted for. In the rare event that all genes have a value of zero, the recombination function will stop looking for genes after 100 failed attempts, so as to prevent an infinite loop.

What portion of mutations are regular versus recombination is controlled by the mutation and recombination chance parameters. In general, through experimentation on test data, it was found a 10% regular mutation chance and 90% recombination mutation chance lead to the best overall performance.

4.5.5 Elitism

Elitism refers to how many of the top-performing chromosomes are kept between generations. If this value is not at least 1, the algorithm has the potential to regress, as it discards the previous top-performing solution. Higher values of elitism will encourage the perpetuation of the most successful genes, potentially improving time to convergence. However, if the value is too high, there is the potential to harm genetic diversity. If half the population are near-identical chromosomes that all perform approximately the same, and thus are kept in the population from generation to generation, genetic diversity will plummet as these near-identical chromosomes only crossover with each other. New genes and gene combinations are needed to encourage a thorough search of the solution space.

4.5.6 Warm Start using GMB Heuristic

Rather than starting from a random initial population, by first solving the GMB heuristic, its solution can be given to the GA as a starting point from which it can improve on. This leads to improved performance compared to either method on its own, which will be explored in more detail in section 5.3.7.

4.5.7 GA Pseudocode

Algorithm 3 contains the pseudocode for the GA.

4.6 Benders Decomposition

BD is a solution method where the problem is solved by fixing the complicating variables, allowing the model to be decomposed into smaller sub-problems and a master problem. These sub-problems are solved to generate an upper bound and add cuts to a master problem. The master problem is then solved to generate a lower bound and new fixed variables for the sub-problem. This repeats until the

Algorithm 3 GA Pseudo Code

Set GA Parameters
Generate N random starting Chromosomes (z)
if Warm Start **then**
    Run GMB heuristic and add BestZ to starting Chromosomes
end if
while $Generation < MaxGeneration$ **and** Average Improvement of previous K generations $> L$ **do**
    Evaluate Fitness of each Chromosome
    Sort Chromosomes by Fitness
    Perform Crossover to create a new generation of N chromosomes
    Add previous top X performing chromosomes to a new generation
    Mutate new generation of chromosomes
end while
Return Top Performing Chromosome

upper and lower bounds converge. Because BD keeps the integrity of the problem intact and we utilize a logic-based BD [27] rather than generalized BD, it will always converge to the optimal solution. This guarantee of optimality is the primary benefit of using BD compared to the other solution methods explored thus far; however, this comes with a trade-off in run time.

The structure of the model would initially suggest that it would lend itself well to this approach, similar to the Lagrangian relaxation approach. By fixing x_{ijvt} to $\bar{x}_{ijvt}$ the model can be decomposed into the SEIR sub-problem, which can be further decomposed by population zone into $|I|$ identical sub-problems, with the supply chain model serving as the basis for the master problem. However, in traditional BD, the purpose of solving the sub-problem is to use the dual variables to find the extreme points and rays and use them to generate cuts for the master problem. Unfortunately, the non-convex nature of the SEIR sub-problem means that there are no dual variables in the classical sense of convex optimization. This means traditional BD cannot be utilized to solve the problem. Instead, a logic-based BD approach is required to generate cuts.

Benders Sub-problem

The sub-problem is based on the SEIR portion of the model and can be further decomposed by population zone into $|I|$ sub-problems. However, as previously mentioned, the sub-problems are still non-convex, meaning the dual variables required for the master problem in traditional BD do not exist. Additionally, similar to the problem with SP2 for the Lagrangian relaxation approach, even with only a single zone the problem quickly becomes intractable with additional groups and an extended time frame. This limits the applicability of the BD method to either small problems, where the sub-problem can still be solved by some other method, such as commercial solvers, or converting it into a homogeneous problem (i.e., one risk

group but multiple independent zones), which can be solved instantly through simulation. Once solved, cuts can be derived from the sub-problems using the solutions themselves. The sub-problem is formulated as follows:

$$\min_{S,E,I,R,NE,z} \quad \sum_{g\in G}\sum_{t\in T} w_{ig} NE_{igt} \tag{4.73}$$

$$\text{s.t} \quad NE_{ig(t+1)} = \tau S_{igt} \sum_{g'\in G} C_{gg'} \frac{I_{ig't}}{P_{ig'}} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.74}$$

$$S_{ig(t+1)} \geq S_{igt} - NE_{ig(t+1)} - \sum_{v\in V} \eta_v z_{igv(t+1)} \quad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.75}$$

$$E_{ig(t+1)} = \left(1 - \frac{1}{\alpha_1}\right) E_{igt} + NE_{ig(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.76}$$

$$I_{ig(t+1)} = \left(1 - \frac{1}{\alpha_2}\right) I_{igt} + \frac{E_{igt}}{\alpha_1} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.77}$$

$$R_{ig(t+1)} = R_{igt} + \frac{I_{igt}}{\alpha_2} + \sum_{v\in V} \eta_v z_{igv(t+1)} \qquad \forall i \in I, g \in G, t \in T\backslash\{t_n\} \tag{4.78}$$

$$\sum_{g\in G} z_{igvt} \leq b_v \sum_{j\in J} \overline{x}_{ijvt} \qquad \forall i \in I, v \in V, t \in T \tag{4.79}$$

$$\sum_{g\in G}\sum_{v\in V} z_{igvt} \leq a_i \qquad \forall i \in I, t \in T \tag{4.80}$$

$$z_{igvt}, S_{igt}, E_{igt}, I_{igt}, R_{ijt}, NE_{igt} \geq 0 \qquad \forall g \in G, t \in T \tag{4.81}$$

Benders Master Problem

The Benders master problem is based on the supply chain portion of the model, formulated below, where θ_i is an approximation of the objective function of the sub-problem i. By solving the master problem with cuts generated from the sub-problems (Θ_i), new values for x_{ijvt} can be identified and passed to the sub-problems. Through this process of continuously generating new cuts and new solutions, eventually, the optimal solution will be found. The Benders master problem is formulated as follows:

$$\min_{x,y,r} \quad \sum_{i\in I}\theta_i + \beta\Big(\sum_{i\in I}\sum_{j\in J}\sum_{v\in V}\sum_{t\in T} c_{ijv}x_{ijvt} + \sum_{j\in J}\sum_{v\in V}\sum_{t\in T} d_{jv}y_{jvt}\Big) \tag{4.82}$$

$$\text{s.t} \quad \sum_{j\in J} y_{jvt} \le q_{vt} \qquad \forall v\in V, t\in T \tag{4.83}$$

$$\sum_{i\in I}\sum_{v\in V} \kappa_v x_{ijvt} \le p_j \qquad \forall j\in J, t\in T \tag{4.84}$$

$$r_{jvt} - r_{jv(t-1)} = y_{jvt} - \sum_{i\in I} x_{ijvt} \qquad \forall j\in J, v\in V, t\in T \tag{4.85}$$

$$\sum_{v\in V} \kappa_v r_{jvt} \le s_j \qquad \forall j\in J, t\in T \tag{4.86}$$

$$r_{jvt} = r^0_{jv} \qquad \forall j\in J, v\in V \tag{4.87}$$

$$\theta_i \in \Theta_i \qquad \forall i\in I \tag{4.88}$$

$$x_{ijvt}, y_{jvt}, r_{jvt}, \theta_i \ge 0 \qquad \forall i\in I, j\in J, v\in V, t\in T \tag{4.89}$$

$$x_{ijvt}, y_{jvt}, r_{jvt} \in \mathbb{Z} \tag{4.90}$$

Generating Logic-Based Benders Cuts

It can be easily proven that the optimal value of Benders sub-problem i cannot be improved by sending fewer lots of vaccines to zone i in at least some periods while keeping the shipped amounts constant in the remaining periods. The sub-problem's

objective function seeks to minimize total new exposures. New exposures increase as a function of the susceptible population. By administering more vaccines we can only ever lower the susceptible population, thus lowering new exposures. If we know the previous allocation was optimal, and fewer vaccines are sent to a zone, total new exposures thus necessarily will either have to be higher than the previous allocation or equal, if the vaccines had no impact. Thus sending fewer vaccines might lead to an equivalent solution, but never an improvement. This property can be used to obtain lower bounds on the optimal values of the sub-problems, and thus generate Benders optimality cuts. Let x^m_{ijvt} and x^{m+1}_{ijvt} be partial optimal solutions to the master problem (i.e., the number of vaccine lots v sent to zone i from DC j in period t) in iterations m and $m+1$, respectively. If $\sum_{j\in J} x^{m+1}_{ijvt} \leq \sum_{j\in J} x^m_{ijvt} \forall v \in V, t \in T$, then $\sum_{g\in G}\sum_{t\in T} w_{ig} NE^m_{igt} \leq \sum_{g\in G}\sum_{t\in T} w_{ig} NE^{m+1}_{igt}$, meaning that $\sum_{g\in G}\sum_{t\in T} w_{ig} NE^m_{igt}$ servers as a lower bound on the optimal value of SP_i in the next iteration. This lower bound can be used to generate the Benders optimality cut:

$$\theta_i \geq \sum_{g\in G}\sum_{t\in T} w_{ig} NE^m_{igt}, \quad \text{if and only if} \quad \sum_{j\in J} x_{ijvt} \leq \sum_{j\in J} x^m_{ijvt}, \quad \forall v \in V, t \in T \tag{4.91}$$

which can be linearized by introducing the binary variables o^m_i and u^m_{ivt} and the following set of constraints:

$$\theta_i \geq o^m_i \sum_{g\in G}\sum_{t\in T} w_{ig} NE^m_{igt} \qquad \forall i \in I, m \in M \tag{4.92}$$

$$\sum_{j\in J} x^m_{ijvt} - \sum_{j\in J} x_{ijvt} + 1 \leq M u^m_{ivt} \qquad \forall i \in I, v \in V, t \in T, m \in M \tag{4.93}$$

$$o^m_i \geq \frac{\sum_{v\in V}\sum_{t\in T} u^m_{ivt} + 1}{|V|.|T|} - 1 \qquad \forall i \in I, m \in M \tag{4.94}$$

$$o^m_i, u^m_{ivt} \in \{0, 1\} \tag{4.95}$$

The first constraint (4.92) ensures that if o_i^m has a value of 1, the lower bound on θ_i is NE_{igt}^m, otherwise the lower bound is 0. The second constraint (4.93) forces the binary variable u_{ivt}^m to take a value of 1 if the number of vaccine lots (of type v in period t) sent to zone i in iteration $m+1$ is less than or equal to those sent in iteration m. If, instead, more lots of vaccines are sent, u_{ivt}^m can take on a value of 0 or 1. The last constraint (4.94) forces o_i^m to take the value 1 if and only if all $u_{ivt}^m = 1$. Note that if given the option of choosing 0 or 1 by constraint (4.93), the decision variable u_{ivt}^m will prefer to take on a value of 0, as this allows o_i^m to be 0, which in turn allows the lower bound on θ_i to be relaxed (4.92). This allows the model to ignore the epidemiological impact of a region in the objective function so long as it sends more lots to that zone than in previous iterations.

Initially, the entire epidemiological impact can be ignored in the master problem by sending one additional lot to each zone in each iteration. Over time, the number of lots sent to a zone will gradually increase until there is no more supply to draw from, after which the model will begin sending different combinations of lots, allowing it to ignore only a portion of the zones each iteration, but not the entire epidemiological impact. Once all possible combinations that allow the model to send more lots of vaccines to a zone than in any previous iteration have been exhausted, the model can be solved instantly. All the cuts generated by the sub-problems collectively allow the model to evaluate the benefit gained from any potential allocation of vaccine lots, and it chooses the best-performing allocation among them.

BD Pseudocode

Algorithm 4 contains the pseudocode for the logic-based BD algorithm.

Algorithm 4 BD Pseudocode

UB = inf, LB = -inf
while $UB - LB > \epsilon$ **do**
    for $i = 1 : I$ **do**
        Solve Benders sub-problem for each zone using $\overline{x}_{ijvt}$
        Extract values of NE^m_{igt} and x^m_{ijvt}
    end for
    Update UB
    Generate and add new cuts (4.92)-(4.94) to Benders Master Problem
    Solve Benders Master problem
    Update LB
    Extract $\overline{x}_{ijvt}$ for use in sub-problem
end while

Chapter 5

Solution Method Comparison

This section compares the performance of the different solution methods described in Chapter 4. The most promising of these methods, i.e., those that could generate feasible solutions in a reasonable time frame, are run on a set of test data for problems of different sizes to evaluate and compare their performance. Because this is only test data, the actual solutions to these problems will not be of interest, instead only how each solution method performs relative to other methods matters. The performance will be judged both in terms of their objective values and run times. If two methods have comparable solutions in terms of their objective values (i.e., within 5% of each other), but one was able to arrive at its solution faster, then it is considered preferable over the other.

Not all solution methods covered in Chapter 4 worked as intended, in the sense that the solutions they produced were of poor quality, infeasible to the original formulation, or had prohibitively long solve times with no guarantee of optimality. Even though there are no numerical results for such solutions methods, they will still be covered alongside the others in the method comparison section 5.3, with a justification as to why they did not work and potential future paths to follow.

The solution methods for which numerical results could be obtained are the Mountain Climbing heuristic (MC), the Greedy Marginal Benefit heuristic (GMB), the Genetic Algorithm (GA) and the GA warm started with the GMB heuristic (GA+GMB), and Benders Decomposition (BD). In the case of the GMB heuristic, the base ten proportioning scheme was used, as described in section 4.4.1. In the case of the methods that use the GA, the stopping condition is set to be after 1000

generations or if less than 0.01 average improvement per generation is seen in the previous 100 generations. In the case of BD, due to the restrictions on the problem size that BD can be used solved in a reasonable time, it is compared on a separate set of problems to the other methods. It could only be used on a small set of scenarios where it can arrive at the optimal solution in less than 24 hours.

In addition, performance is evaluated for two potential vaccination policies: 1) Non-Intervention (NI), i.e., sending no vaccines, and 2) Pro-rata allocation (PR), i.e., allocating supply proportional to population size. Evaluating the scenario of Non-Intervention allows for establishing a baseline, worst-case scenario from which the impact of other strategies can be judged fairly. For example, say the objective values found using two solution methods differ by 1,000. If the difference from NI is only 5,000, that represents a significant difference in quality between the two solution methods. If, instead, the difference from NI is 1,000,000, the two methods actually performed comparably, especially as small changes to the vaccination strategy will naturally propagate over time. Every potential vaccination strategy will fall somewhere between this NI strategy and the true optimal allocation strategy.

Meanwhile, Pro-rata allocation was chosen as it is a very common approach to allocating vaccines because it does not end up favoring any one group over another. There are actually two PR approaches used for comparison in this section. The first is a rigid application of Pro-rata, allocating vaccines strictly proportional to the population of a zone and group, regardless of whether they can actually use those vaccines or not. This leads to two problems: 1) it tends to lead to wasting vaccines once a portion of the population has been fully immunized. 2) a purely Pro-rata allocation of vaccines might be infeasible to the original problem and violate administrative or logistical capacity constraints. Therefore, a corrected version of the PR strategy (PR*) was developed, where if a population group has been fully immunized, those vaccines are no longer allocated to them, saving costs, and if a region exceeds its administrative capacity, the number of vaccines sent to the region

is reduced to the administrative capacity. The performance of a simple rule-of-thumb vaccine allocation policy such as PR is perhaps a more fair comparison than Non-Intervention, as in reality, it would be unrealistic, or at least irresponsible, to do nothing in the face of an epidemic if vaccination was a possibility. If the vaccination strategies identified by the solution methods cannot significantly improve on a simple strategy like proportionally allocating available vaccines, then they likely are not worth the increased complexity of gathering and verifying the data, and building, validating, and running the algorithms.

5.1 Data Generation

Pseudo-random test data was generated specifically for use in these tests, though the values are based on those that might be experienced in a real scenario. Population zones were placed by generating random uniformly distributed x and y coordinates between 0 and 1. The overall population of each zone is a random uniformly distributed integer number between 100,000 and 1,000,000, which is then randomly divided among 5 risk groups. The initial infected population for each group is a random integer number between 1 and 1000, and the initial exposed population is twice that number. The initial susceptible population is the group's overall population minus the infected and exposed populations, and finally, the initial removed population is set to 0.

To create the contact matrix, a random number between 0 and 1 is generated for each combination of groups. These numbers are then divided by their sum across the group to normalize them and get what percentage of a group's contacts are with members of each group per time period. A random number between 15 and 40 is generated to represent each group's total contacts per week. The total contacts are then multiplied by each group's percentage contact rate for each group, to get the contacts per group per week, which gives the contact matrix.

Two potential vaccines were used: one high efficacy, high cost, and bulky, and one low efficacy, low cost, and non-bulky. The supply of each was kept consistent across all time periods, at 200 and 100 lots of each type, respectively. DCs were placed at the center and each corner of the population zone grid. The center DC had increased capacity and slightly decreased vaccine acquisition cost compared to the corner DCs. The distances between the DCs and population zones are measured as the Euclidean distance on a 1-by-1 grid, where 1 unit is equivalent to 300 km.

The transmission rate was set to 5% ($\tau = 0.05$), the average exposure duration and average infection duration were both considered to be two time periods ($\alpha_1 = 2$, $\alpha_2 = 2$), and cost per km is set at \$2.00 per 20,000 cubic inches of product shipped. These values are all tuned to represent a fast-spreading disease, but one that would still be progressing after 20-30 weeks. A more infectious disease is better for these tests as it allows differences between vaccination strategies to be accentuated. At the same time, we do not want the disease to be too infectious, as then the vaccines would not have an opportunity to take effect before everyone became infected. As we are interested in preventing exposures with this test data, the weight parameters (w_{ig}) are simply set to 1. A value of 0.001 was chosen for β; representing that preventing 1 additional exposure was worth spending an additional \$1,000, as it was sufficiently low to encourage the model to prioritize vaccination, but high enough that the costs were still relevant to the objective function.

5.2 Numerical Results

This section will present the results of the different solution methods for problems of varying sizes and levels of vaccine supply. Problem size, (i.e., how many regions, DCs, groups, vaccine types, and time periods are used in the model), and the available supply of vaccines are the primary factors that impact both solution quality and run time. Parameters other than the supply of vaccines would have little to no impact

on run time, as they have no impact on how the solving algorithms behave. One method may perform slightly better or worse than it typically does under specific parameter settings, but in general, the relative performance of the objective values of different methods should be comparable across different scenarios of the same-sized problem.

Table 5.1 depicts the objective values obtained using NI, PR, PR*, MC, GA, GMB, and GA+GMB for different problem sizes. Table 5.2 shows the time to solution for each method, other than NI, PR, and PR*, as they can be solved for almost instantaneously. Table 5.3 contains the relative improvement of each solution method measured as the percentage improvement over the NI scenario. As previously mentioned, the solutions from the PR category are usually either infeasible or wasteful. Therefore, rather than comparing to PR, Table 5.4 contains the relative improvement of each solution method measured as a percentage improvement over using a PR* allocation policy. Table 5.5 contains the relative improvement of the MC and GA+GMB solution methods measured as a percentage improvement over GMB. Table 5.6 contains the results for a set of problems for which GUROBI can achieve a relatively small optimality gap within a reasonable time frame. The table contains the objective values obtained from using the acrshortga+GMB solution method, and the upper (UB) and lower (LB) bounds obtained from GUROBI. Alongside these values is the gap between the UB and LB, the gap between GA+GMB and the LB (i.e., the maximum potential optimality gap of the solution method), and the gap between GA+GMB and the UB (i.e., the minimum potential optimality gap of the solution method). Table 5.7 contains the objective values obtained from using NI, PR*, GA+GMB, and BD on a separate set of problems that are tailored to allow the BD method to solve in a reasonable time frame. When the supply column contains two numbers (e.g., 200 100), the two numbers represent the supply of two different types of vaccines.

For Tables 5.3, 5.4, and 5.5, the relative improvement is calculated using equation

(5.1).

$$\% \text{ Improvement from } X = \frac{\text{Obj } X - \text{Obj } Y}{\text{Obj } X}. \tag{5.1}$$

Table 5.1: Best Objective Value obtained using different policies or solution methods

ID	I	J	G	V	T	Supply	NI	PR	PR*	MC	GA	GMB	GA+GMB
1	1	1	3	1	10	200	81,225	77,560	39,518	31,331	31,692	31,380	31,283
2	1	1	3	1	20	200	226,696	161,615	46,594	35,803	38,050	36,015	35,777
3	1	1	3	1	30	200	230,285	245,671	46,630	35,551	37,958	36,021	35,783
4	1	1	5	1	10	200	193,428	79,228	115,721	101,289	102,537	113,709	101,820
5	1	1	5	1	20	200	355,756	163,284	154,069	134,161	141,507	136,291	136,030
6	1	1	5	1	30	200	356,175	247,339	154,218	134,161	140,186	136,417	136,153
7	5	1	1	1	10	200	27,427	78,898	24,523	17,270	18,779	17,311	17,269
8	5	1	1	1	20	200	81,315	162,943	24,729	25,024	23,101	21,514	21,472
9	5	1	1	1	30	200	155,797	246,989	24,736	25,070	24,235	21,963	21,871
10	5	1	3	1	10	200	642,859	165,285	304,170	254,563	263,745	261,269	259,936
11	5	1	3	1	20	200	200,3671	249,328	359,469	290,432	313,687	303,236	301,358
12	5	1	3	1	30	200	2,052,450	333,372	360,049	290,435	317,695	302,637	301,798
13	5	1	5	1	10	200	1,848,518	520,979	1,116,939	1,039,158	9,76,538	1,007,266	966,211
14	5	1	5	1	20	200	3,337,877	605,023	1,451,301	1,403,106	1,333,485	1,288,399	1,282,855
15	5	1	5	1	30	200	3,342,509	689,067	1,452,694	1,403,195	1,334,182	1,283,366	1,280,651
16	10	1	1	1	10	200	57,181	86,229	48,843	35,282	37,097	35,301	35,212
17	10	1	1	1	20	200	167,061	170,267	49,376	68,927	41,967	40,436	40,319
18	10	1	1	1	30	200	313,326	254,304	49,391	125,906	45,295	40,901	40,823
19	10	1	3	1	10	200	1,332,858	483,495	810,452	521,013	536,814	524,586	520,978
20	10	1	3	1	20	200	3,476,941	567,880	1,117,495	578,511	637,481	602,424	592,424
21	10	1	3	1	30	200	3,543,272	651,917	1,119,679	578,140	649,008	600,627	592,472
22	10	1	5	1	10	200	3,510,689	1,758,471	2,530,597	1,950,506	1,885,165	1,902,153	1,839,876
23	10	1	5	1	20	200	5,822,143	2,002,176	3,509,954	2,485,905	2,520,173	2,364,282	2,350,566
24	10	1	5	1	30	200	5,828,652	2,086,213	3,513,359	2,495,533	2,628,966	2,359,958	2,349,315
25	10	5	5	1	10	200	3,510,689	1,758,460	2,530,592	1,936,633	1,881,059	1,902,153	1,840,482
26	10	5	5	1	20	200	5,822,143	2,002,152	3,509,946	2,495,358	2,536,545	2,364,282	2,349,713
27	10	5	5	1	30	200	5,828,652	2,086,177	3,513,350	2,495,546	2,613,861	2,359,958	2,348,927
28	20	1	5	1	10	200	6,781,450	4,708,816	4,707,054	4,281,687	4,122,334	4,074,798	3,924,889
29	20	1	5	1	20	200	9,906,374	5,751,306	5,780,962	5,497,124	5,846,025	4,929,931	4,898,895
30	20	1	5	1	30	200	9,915,241	5,835,320	5,785,310	5,525,863	6,664,304	4,965,128	4,932,607
31	20	5	5	1	10	200	6,781,450	4,708,814	4,707,053	4,243,462	4,127,983	4,074,798	3,923,061
32	20	5	5	1	20	200	9,906,374	5,751,303	5,780,960	5,541,834	5,916,761	4,929,931	4,898,124
33	20	5	5	1	30	200	9,915,241	5,835,315	5,785,309	5,554,755	5,885,054	4,929,929	4,900,590
34	10	1	5	2	10	200 100	3,510,689	1,491,230	2,283,716	1,950,506	2,028,398	1,952,150	1,839,800
35	10	1	5	2	20	200 100	5,822,143	1,639,146	2,948,207	2,495,063	2,922,411	2,364,279	2,350,164
36	10	1	5	2	30	200 100	5,828,652	1,472,262	2,979,429	2,496,049	3,218,710	2,359,955	2,348,299
37	10	1	5	1	20	10	5,822,143	5,551,445	5,551,443	5,526,520	5,515,449	5,454,081	5,450,360
38	10	1	5	1	20	50	5,822,143	4,569,989	4,580,633	4,487,355	4,363,927	4,247,665	4,241,449
39	10	1	5	1	20	100	5,822,143	3,494,181	3,509,954	3,535,909	3,223,271	3,143,497	3,124,051
40	10	1	5	1	20	150	5,822,143	2,627,874	2,649,778	2,502,231	2,518,972	2,364,051	2,350,985

Table 5.2: CPU times for different solution methods

ID	I	J	G	V	T	Supply	MC	GA	GMB	GA+GMB
1	1	1	3	1	10	200	0:00:00	0:01:16	0:00:02	0:00:40
2	1	1	3	1	20	200	0:00:00	0:03:17	0:00:09	0:01:12
3	1	1	3	1	30	200	0:00:04	0:05:13	0:00:22	0:01:41
4	1	1	5	1	10	200	0:00:00	0:00:06	0:02:23	0:03:18
5	1	1	5	1	20	200	0:00:07	0:08:00	0:00:26	0:03:42
6	1	1	5	1	30	200	0:00:04	0:16:59	0:00:59	0:04:20
7	5	1	1	1	10	200	0:00:00	0:03:02	0:00:10	0:01:08
8	5	1	1	1	20	200	0:00:05	0:14:36	0:00:44	0:02:17
9	5	1	1	1	30	200	0:02:06	0:30:09	0:01:39	0:04:18
10	5	1	3	1	10	200	0:00:03	0:25:49	0:00:47	0:05:31
11	5	1	3	1	20	200	0:00:16	0:45:38	0:03:26	0:25:34
12	5	1	3	1	30	200	0:00:30	1:07:02	0:07:57	0:31:58
13	5	1	5	1	10	200	0:00:18	0:36:36	0:02:03	0:17:35
14	5	1	5	1	20	200	0:01:13	1:21:41	0:10:53	1:35:56
15	5	1	5	1	30	200	0:03:10	1:54:03	0:23:42	2:16:10
16	10	1	1	1	10	200	0:00:00	0:15:08	0:00:41	0:05:54
17	10	1	1	1	20	200	0:00:11	0:42:27	0:02:45	0:15:21
18	10	1	1	1	30	200	0:02:35	1:04:48	0:06:34	0:13:50
19	10	1	3	1	10	200	0:00:00	0:43:08	0:02:52	0:31:19
20	10	1	3	1	20	200	0:00:08	1:25:54	0:12:48	1:38:26
21	10	1	3	1	30	200	0:00:45	2:10:19	0:29:50	2:39:26
22	10	1	5	1	10	200	0:00:03	1:12:41	0:08:12	1:21:55
23	10	1	5	1	20	200	0:00:06	2:34:43	0:38:59	3:44:40
24	10	1	5	1	30	200	0:04:08	3:37:58	1:26:56	5:02:51
25	10	5	5	1	10	200	0:00:03	1:18:07	0:08:07	0:43:16
26	10	5	5	1	20	200	0:00:53	2:34:59	0:42:15	3:53:18
27	10	5	5	1	30	200	0:00:12	3:44:20	1:29:12	5:15:54
28	20	1	5	1	10	200	0:00:23	2:16:25	0:31:08	2:47:21
29	20	1	5	1	20	200	0:21:11	4:43:00	2:26:52	9:34:39
30	20	1	5	1	30	200	0:36:59	7:21:22	5:44:37	12:58:20
31	20	5	5	1	10	200	0:00:56	2:25:23	0:32:43	2:56:44
32	20	5	5	1	20	200	0:04:45	4:54:52	2:30:43	9:56:40
33	20	5	5	1	30	200	0:07:16	5:05:04	2:35:06	5:00:48
34	10	1	5	2	10	200 100	0:00:34	1:24:45	0:17:12	1:17:43
35	10	1	5	2	20	200 100	0:18:29	2:50:23	1:16:35	5:24:46
36	10	1	5	2	30	200 100	0:50:10	4:14:14	3:05:07	8:07:46
37	10	1	5	1	20	10	0:00:05	2:28:37	0:28:03	1:49:36
38	10	1	5	1	20	50	0:01:11	2:29:49	0:35:22	3:02:23
39	10	1	5	1	20	100	0:03:37	2:28:05	0:36:02	3:02:58
40	10	1	5	1	20	150	0:00:48	2:29:52	0:37:18	3:08:00

Table 5.3: Percent Improvement from NI

ID	I	J	G	V	T	Supply	PR*	MC	GA	GMB	GA+GMB
1	1	1	3	1	10	200	51.35%	61.43%	60.98%	61.37%	**61.49%**
2	1	1	3	1	20	200	79.45%	84.21%	83.22%	84.11%	**84.22%**
3	1	1	3	1	30	200	79.75%	**84.56%**	83.52%	84.36%	84.46%
4	1	1	5	1	10	200	40.17%	**47.63%**	46.99%	41.21%	47.36%
5	1	1	5	1	20	200	56.69%	**62.29%**	60.22%	61.69%	61.76%
6	1	1	5	1	30	200	56.70%	**62.33%**	60.64%	61.70%	61.77%
7	5	1	1	1	10	200	10.59%	37.03%	31.53%	36.88%	**37.04%**
8	5	1	1	1	20	200	69.59%	69.23%	71.59%	73.54%	**73.59%**
9	5	1	1	1	30	200	84.12%	83.91%	84.44%	85.90%	**85.96%**
10	5	1	3	1	10	200	52.68%	**60.40%**	58.97%	59.36%	59.57%
11	5	1	3	1	20	200	82.06%	**85.50%**	84.34%	84.87%	84.96%
12	5	1	3	1	30	200	82.46%	**85.85%**	84.52%	85.25%	85.30%
13	5	1	5	1	10	200	39.58%	43.78%	47.17%	45.51%	**47.73%**
14	5	1	5	1	20	200	56.52%	57.96%	60.05%	61.40%	**61.57%**
15	5	1	5	1	30	200	56.54%	58.02%	60.08%	61.60%	**61.69%**
16	10	1	1	1	10	200	14.58%	38.30%	35.12%	38.26%	**38.42%**
17	10	1	1	1	20	200	70.44%	58.74%	74.88%	75.80%	**75.87%**
18	10	1	1	1	30	200	84.24%	59.82%	85.54%	86.95%	**86.97%**
19	10	1	3	1	10	200	39.19%	**60.91%**	59.72%	60.64%	**60.91%**
20	10	1	3	1	20	200	67.86%	**83.36%**	81.67%	82.67%	82.96%
21	10	1	3	1	30	200	68.40%	**83.68%**	81.68%	83.05%	83.28%
22	10	1	5	1	10	200	27.92%	44.44%	46.30%	45.82%	**47.59%**
23	10	1	5	1	20	200	39.71%	57.30%	56.71%	59.39%	**59.63%**
24	10	1	5	1	30	200	39.72%	57.19%	54.90%	59.51%	**59.69%**
25	10	5	5	1	10	200	27.92%	44.84%	46.42%	45.82%	**47.57%**
26	10	5	5	1	20	200	39.71%	57.14%	56.43%	59.39%	**59.64%**
27	10	5	5	1	30	200	39.72%	57.18%	55.15%	59.51%	**59.70%**
28	20	1	5	1	10	200	30.59%	36.86%	39.21%	39.91%	**42.12%**
29	20	1	5	1	20	200	41.64%	44.51%	40.99%	50.23%	**50.55%**
30	20	1	5	1	30	200	41.65%	44.27%	32.79%	49.92%	**50.25%**
31	20	5	5	1	10	200	30.59%	37.43%	39.13%	39.91%	**42.15%**
32	20	5	5	1	20	200	41.64%	44.06%	40.27%	50.23%	**50.56%**
33	20	5	5	1	30	200	41.65%	43.98%	40.65%	50.28%	**50.58%**
34	10	1	5	2	10	200 100	34.95%	44.44%	42.22%	44.39%	**47.59%**
35	10	1	5	2	20	200 100	49.36%	57.15%	49.81%	59.39%	**59.63%**
36	10	1	5	2	30	200 100	48.88%	57.18%	44.78%	59.51%	**59.71%**
37	10	1	5	1	20	10	4.65%	5.08%	5.27%	6.32%	**6.39%**
38	10	1	5	1	20	50	21.32%	22.93%	25.05%	27.04%	**27.15%**
39	10	1	5	1	20	100	39.71%	39.27%	44.64%	46.01%	**46.34%**
40	10	1	5	1	20	150	54.49%	57.02%	56.73%	59.40%	**59.62%**

Table 5.4: Percent Improvement from PR*

ID	I	J	G	V	T	Supply	MC	GA	GMB	GA+GMB
1	1	1	3	1	10	200	20.72%	19.81%	20.59%	**20.84%**
2	1	1	3	1	20	200	23.16%	18.34%	22.71%	**23.22%**
3	1	1	3	1	30	200	**23.76%**	18.60%	22.75%	23.26%
4	1	1	5	1	10	200	**12.47%**	11.39%	1.74%	12.01%
5	1	1	5	1	20	200	**12.92%**	8.15%	11.54%	11.71%
6	1	1	5	1	30	200	**13.01%**	9.10%	11.54%	11.71%
7	5	1	1	1	10	200	**29.58%**	23.42%	29.41%	**29.58%**
8	5	1	1	1	20	200	-1.19%	6.59%	13.00%	**13.17%**
9	5	1	1	1	30	200	-1.35%	2.02%	11.21%	**11.58%**
10	5	1	3	1	10	200	**16.31%**	13.29%	14.10%	14.54%
11	5	1	3	1	20	200	**19.21%**	12.74%	15.64%	16.17%
12	5	1	3	1	30	200	**19.33%**	11.76%	15.95%	16.18%
13	5	1	5	1	10	200	6.96%	12.57%	9.82%	**13.49%**
14	5	1	5	1	20	200	3.32%	8.12%	11.22%	**11.61%**
15	5	1	5	1	30	200	3.41%	8.16%	11.66%	**11.84%**
16	10	1	1	1	10	200	27.76%	24.05%	27.73%	**27.91%**
17	10	1	1	1	20	200	-39.60%	15.01%	18.11%	**18.34%**
18	10	1	1	1	30	200	-154.91%	8.29%	17.19%	**17.35%**
19	10	1	3	1	10	200	35.71%	33.76%	35.27%	**35.72%**
20	10	1	3	1	20	200	**48.23%**	42.95%	46.09%	46.99%
21	10	1	3	1	30	200	**48.37%**	42.04%	46.36%	47.09%
22	10	1	5	1	10	200	22.92%	25.51%	24.83%	**27.29%**
23	10	1	5	1	20	200	29.18%	28.20%	32.64%	**33.03%**
24	10	1	5	1	30	200	28.97%	25.17%	32.83%	**33.13%**
25	10	5	5	1	10	200	23.47%	25.67%	24.83%	**27.27%**
26	10	5	5	1	20	200	28.91%	27.73%	32.64%	**33.06%**
27	10	5	5	1	30	200	28.97%	25.60%	32.83%	**33.14%**
28	20	1	5	1	10	200	9.04%	12.42%	13.43%	**16.62%**
29	20	1	5	1	20	200	4.91%	-1.13%	14.72%	**15.26%**
30	20	1	5	1	30	200	4.48%	-15.19%	14.18%	**14.74%**
31	20	5	5	1	10	200	9.85%	12.30%	13.43%	**16.66%**
32	20	5	5	1	20	200	4.14%	-2.35%	14.72%	**15.27%**
33	20	5	5	1	30	200	3.99%	-1.72%	14.79%	**15.29%**
34	10	1	5	2	10	200 100	14.59%	11.18%	14.52%	**19.44%**
35	10	1	5	2	20	200 100	15.37%	0.87%	19.81%	**20.28%**
36	10	1	5	2	30	200 100	16.22%	-8.03%	20.79%	**21.18%**
37	10	1	5	1	20	10	0.45%	0.65%	1.75%	**1.82%**
38	10	1	5	1	20	50	2.04%	4.73%	7.27%	**7.40%**
39	10	1	5	1	20	100	-0.74%	8.17%	10.44%	**10.99%**
40	10	1	5	1	20	150	5.57%	4.94%	10.78%	**11.28%**

Table 5.5: Percent Improvement from GMB

ID	I	J	G	V	T	Supply	MC	GA+GMB
1	1	1	3	1	10	200	0.16%	**0.31%**
2	1	1	3	1	20	200	0.59%	**0.66%**
3	1	1	3	1	30	200	**1.31%**	0.66%
4	1	1	5	1	10	200	**10.92%**	10.46%
5	1	1	5	1	20	200	**1.56%**	0.19%
6	1	1	5	1	30	200	**1.65%**	0.19%
7	5	1	1	1	10	200	**0.24%**	**0.24%**
8	5	1	1	1	20	200	-16.32%	**0.20%**
9	5	1	1	1	30	200	-14.15%	**0.42%**
10	5	1	3	1	10	200	**2.57%**	0.51%
11	5	1	3	1	20	200	**4.22%**	0.62%
12	5	1	3	1	30	200	**4.03%**	0.28%
13	5	1	5	1	10	200	-3.17%	**4.08%**
14	5	1	5	1	20	200	-8.90%	**0.43%**
15	5	1	5	1	30	200	-9.34%	**0.21%**
16	10	1	1	1	10	200	0.06%	**0.25%**
17	10	1	1	1	20	200	-70.46%	**0.29%**
18	10	1	1	1	30	200	-207.83%	**0.19%**
19	10	1	3	1	10	200	0.68%	**0.69%**
20	10	1	3	1	20	200	**3.97%**	1.66%
21	10	1	3	1	30	200	**3.74%**	1.36%
22	10	1	5	1	10	200	-2.54%	**3.27%**
23	10	1	5	1	20	200	-5.14%	**0.58%**
24	10	1	5	1	30	200	-5.74%	**0.45%**
25	10	5	5	1	10	200	-1.81%	**3.24%**
26	10	5	5	1	20	200	-5.54%	**0.62%**
27	10	5	5	1	30	200	-5.75%	**0.47%**
28	20	1	5	1	10	200	-5.08%	**3.68%**
29	20	1	5	1	20	200	-11.51%	**0.63%**
30	20	1	5	1	30	200	-11.29%	**0.65%**
31	20	5	5	1	10	200	-4.14%	**3.72%**
32	20	5	5	1	20	200	-12.41%	**0.65%**
33	20	5	5	1	30	200	-12.67%	**0.60%**
34	10	1	5	2	10	200 100	0.08%	**5.76%**
35	10	1	5	2	20	200 100	-5.53%	**0.60%**
36	10	1	5	2	30	200 100	-5.77%	**0.49%**
37	10	1	5	1	20	10	-1.33%	**0.07%**
38	10	1	5	1	20	50	-5.64%	**0.15%**
39	10	1	5	1	20	100	-12.48%	**0.62%**
40	10	1	5	1	20	150	-5.85%	**0.55%**

Table 5.6: GUROBI Results and Optimality Gap

ID	I	J	G	v	T	Supply	GA+GMB	UB	LB	UB to LB Gap	GA+GMB to LB Gap	GA+GMB to UB Gap
1	1	1	5	1	6	200	38,404	38,404	38,404	0.00%	0.00%	0.00%
2	1	1	5	1	9	200	86,135	86,011	74,695	13.20%	13.28%	0.14%
3	3	1	2	1	6	200	23,960	23,679	23,679	0.00%	1.18%	1.18%
4	3	1	2	1	7	200	27,663	27,370	27,370	0.00%	1.06%	1.06%
5	3	1	2	1	8	200	30,701	30,382	29,459	3.04%	4.05%	1.04%
6	3	1	2	1	9	200	33,330	32,776	30,327	7.47%	9.01%	1.66%
7	3	1	2	1	10	200	35,382	34,744	30,338	12.70%	14.25%	1.80%
8	3	1	5	1	6	200	260,679	258,785	254,120	1.80%	2.52%	0.73%
9	5	1	3	1	6	200	141,487	140,072	137,357	1.94%	2.92%	1.00%
10	5	1	5	1	6	200	393,211	390,571	371,422	4.90%	5.54%	0.67%
11	10	1	2	1	6	200	70,918	70,714	69,339	1.94%	2.23%	0.29%
12	10	1	5	1	10	200	1,839,876	2,027,311	701,373	65.40%	61.88%	-10.19%

Table 5.7: BD Results

I	J	G	V	T	Supply	NI	PR*	GA+GMB	Objective	Time	Iterations
2	1	1	1	10	1	10,105	9,743	9,339	9,022	0:03:57	511
3	1	1	1	10	1	18,802	18,293	18,019	18,019	0:18:48	792
4	1	1	1	10	1	22,010	21,796	21,234	21,227	0:47:05	1035
5	1	1	1	10	1	27,427	27,342	26,751	26,324	1:33:44	1233
2	1	1	1	5	2	3,630	3,745	3,603	3,603	0:00:07	82
2	1	1	1	6	2	4,771	4,853	4,651	4,651	0:01:55	244
2	1	1	1	7	2	5,988	5,985	5,694	5,694	0:34:33	732

5.2.1 Numerical Results Summary

As can be observed in Table 5.3, all the tested solution methods were able to significantly improve the outcome of a scenario versus doing nothing in the same scenario. In general, this improvement ranged between 40-80%, with the largest differences occurring in problems with fewer zones and groups but more time periods. The impact of additional time periods is obvious, as the more time periods considered, the longer the disease is potentially allowed to propagate in the population. Without intervention, each additional time period will simply add additional opportunities for a person to become infected, until the disease has run its course. If instead some portion of the population is vaccinated, for every additional time period considered, additional infections that would have occurred in the case of Non-Intervention are prevented, as the now vaccinated portion of the population is no longer a vector

for the disease. For example, comparing Non-Intervention against vaccinating 1000 people in the first time period, it will be observed that the outcomes drift further and further apart with each additional period until the disease has run its course. Why is the gap larger for the problems with fewer regions and groups? The most likely reason is simply that the supply relative to the population size shrinks with additional zones and groups, meaning a smaller and smaller fraction of the population can be vaccinated. This hypothesis is supported by the improvement observed in tests 37-40, where the problem size is kept consistent but supply is varied. To properly test this hypothesis, the supply could be kept proportional to the overall population size for each problem. However, as previously mentioned, the point of these tests is not to compare the same method across the set of problems, but to compare the set of methods across the same problem. The fact that vaccines become scarcer for larger problems does not impact that goal.

As previously mentioned, the corrected Pro-rata strategy gives a more realistic point of comparison, compared to the unlikely scenario of doing nothing. From Table 5.4, we generally see an 11-40% improvement using the strategies derived from our tested solution methods when compared to the corrected Pro-rata allocation strategy. This represents a significant improvement, and in real-world scenarios, would support the use of these solution methods over the Pro-rata strategy. From Table 5.4, it becomes obvious that the GA alone was unable to keep up with the other solution methods, as it was never able to outperform any of the other solution methods within the allotted 1000 generations. While in theory the GA could have been left to run for longer to try to improve its objective value, comparing the run times in Table 5.2, it becomes apparent that this improvement would not have been worth the additional run time. Compared to the GMB or MC heuristics, the GA will always take longer to reach a comparable objective value and there is no easy method to predict how long it will take to do so. This is the benefit of warm starting the GA with a solution from the GMB heuristic. It can start with a good solution

in a fraction of the time and improve from there.

As can be seen in Table 5.5, the MC and GMB heuristics were the best-performing individual solutions. The MC heuristic performed better on smaller problems, usually less than 4% better compared to the GMB alone. Meanwhile, the GMB performed better on medium to large problems, usually by 3-12%. On top of this, the MC heuristic is significantly faster in comparison to the GMB heuristic, usually only taking seconds to run whereas the GMB heuristic might take minutes, or running in minutes when the GMB might take hours. This makes it a promising candidate for use in generating a warm start for the GA. The warm start may not be as good as the one from the GMB heuristic, but it could be found in a fraction of the time, with the difference in quality potentially being made up for by the GA.

When given a warm start, the GA was always able to improve on the solution of the GMB heuristic. However, as seen in Table 5.5, this improvement was less than 1% for the majority of the test problems, though in some circumstances it was higher. Does such a small improvement merit several hours of additional run time? Such a small difference is almost certainly within the margins of error of the models' predictions. At the same time, it can be seen that for some of these scenarios, there was potential for improvement of at least 2-4% higher, as the MC heuristic was able to achieve such a solution. Figures 5.1 and 5.2, and Table 5.8 depict an example of the allocation pattern for both strategies where the MC heuristic outperformed the GMB heuristic by approximately 4%.

Figure 5.1: MC Vaccine Allocation

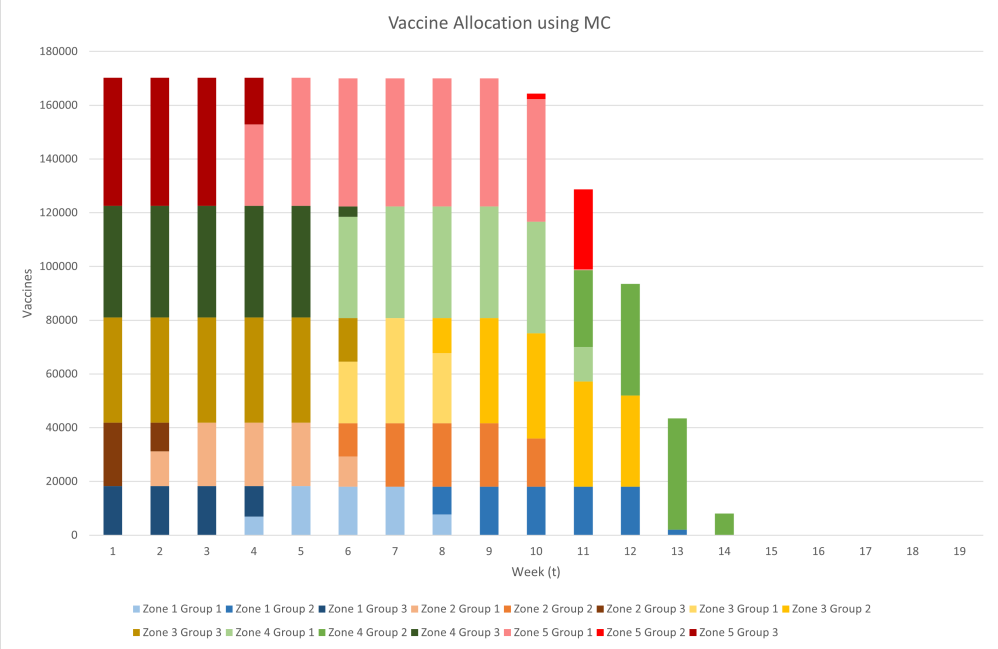


Figure 5.2: GMB Vaccine Allocation

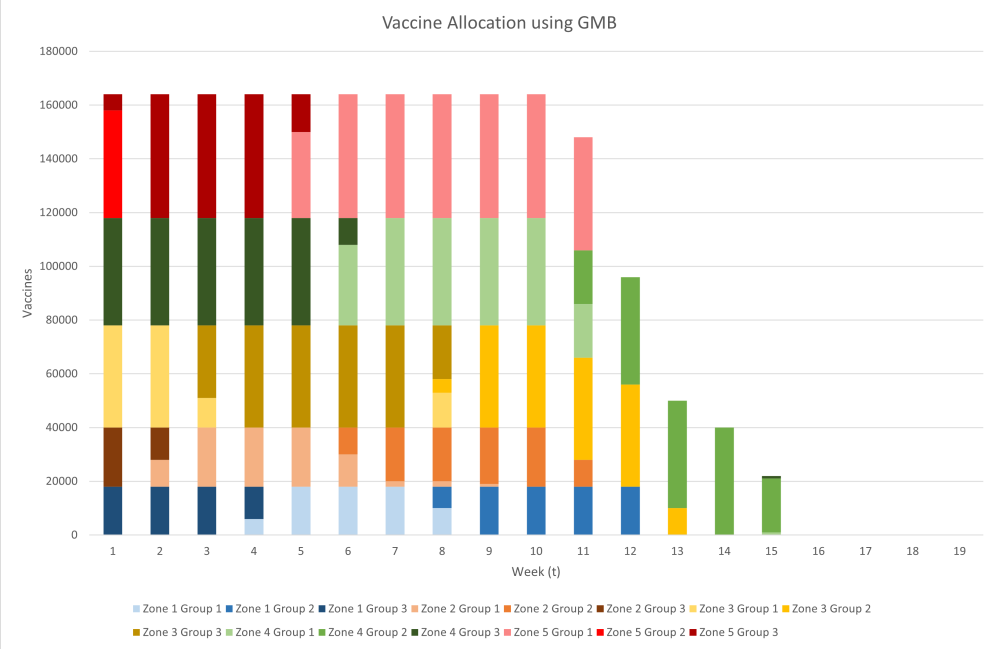


Table 5.8: Allocation Strategies using MC and GMB Heuristics

	MC			GMB		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
t	Zone 1					
1	-	-	18,224	-	-	18,000
2	-	-	18,224	-	-	18,000
3	-	-	18,224	-	-	18,000
4	6,884	-	11,340	6,000	-	12,000
5	18,224	-	-	18,000	-	-
6	18,000	-	-	18,000	-	-
7	18,000	-	-	18,000	-	-
8	7,672	10,328	0	10,000	8,000	-
9	-	18,000	-	-	18,000	-
10	-	18,000	-	-	18,000	-
11	-	18,000	-	-	18,000	-
12	-	18,000	-	-	18,000	-
13	-	2,000	-	-	-	-
14	-	-	-	-	-	-
15	-	-	-	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-
18	-	-	-	-	-	-
19	-	-	-	-	-	-
T	Zone 2					
1	-	-	23,678	-	-	22,000
2	13,021	-	10,657	10,000	-	12,000
3	23,678	-	-	22,000	-	-
4	23,678	-	-	22,000	-	-
5	23,678	-	-	22,000	-	-
6	11,311	12,367	-	12,000	10,000	-
7	-	23,678	-	2,000	20,000	-
8	-	23,678	-	2,000	20,000	-

Table 5.8 – continued from previous page

	MC			GMB		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
9	-	23,678	-	1,000	21,000	-
10	-	18,000	-	-	22,000	-
11	-	-	-	-	10,000	-
12	-	-	-	-	-	-
13	-	-	-	-	-	-
14	-	-	-	-	-	-
15	-	-	-	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-
18	-	-	-	-	-	-
19	-	-	-	-	-	-
T	Zone 3					
1	-	-	39,171	38,000	-	-
2	-	-	39,171	38,000	-	-
3	-	-	39,171	11,000	-	27,000
4	-	-	39,171	-	-	38,000
5	-	-	39,171	-	-	38,000
6	22,870	-	16,301	-	-	38,000
7	39,171	-	-	-	-	38,000
8	26,034	13,137	-	13,000	5,000	20,000
9	-	39,171	-	-	38,000	-
10	-	39,171	-	-	38,000	-
11	-	39,171	-	-	38,000	-
12	-	34,000	-	-	38,000	-
13	-	-	-	-	10,000	-
14	-	-	-	-	-	-
15	-	-	-	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-
18	-	-	-	-	-	-

Table 5.8 – continued from previous page

	MC			GMB		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
19	-	-	-	-	-	-
T	Zone 4					
1	-	-	41,498	-	-	40,000
2	-	-	41,498	-	-	40,000
3	-	-	41,498	-	-	40,000
4	-	-	41,498	-	-	40,000
5	-	-	41,498	-	-	40,000
6	37,620	-	3,878	30,000	-	10,000
7	41,498	-	-	40,000	-	-
8	41,498	-	-	40,000	-	-
9	41,498	-	-	40,000	-	-
10	41,498	-	-	40,000	-	-
11	12,893	28,605	-	20,000	20,000	-
12	-	41,498	-	-	40,000	-
13	-	41,498	-	-	40,000	-
14	-	8,000	-	-	40,000	-
15	-	-	-	1,000	20,000	1,000
16	-	-	-	-	-	-
17	-	-	-	-	-	-
18	-	-	-	-	-	-
19	-	-	-	-	-	-
T	Zone 5					
1	-	-	47,640	-	40,000	6,000
2	-	-	47,640	-	-	46,000
3	-	-	47,640	-	-	46,000
4	30,307	-	17,333	-	-	46,000
5	47640	-	-	32,000	-	14,000
6	47,640	-	-	46,000	-	-
7	47,640	-	-	46,000	-	-
8	47,640	-	-	46,000	-	-

Table 5.8 – continued from previous page

	MC			GMB		
	Group 1	Group 2	Group 3	Group 1	Group 2	Group 3
9	47,640	-	-	46,000	-	-
10	45,664	1,976	-	46,000	-	-
11	241	29,759	-	42,000	-	-
12	-	-	-	-	-	-
13	-	-	-	-	-	-
14	-	-	-	-	-	-
15	-	-	-	-	-	-
16	-	-	-	-	-	-
17	-	-	-	-	-	-
18	-	-	-	-	-	-
19	-	-	-	-	-	-

The overall strategy of both methods can be seen to be very similar. However, the GMB heuristic strategy is allocated entirely in units of 1000, while the MC heuristic is more flexible. It can also be observed that the MC heuristic will generally allocate up to the administrative capacity of a zone when possible, whereas the GMB heuristic will stop just before reaching that point. This leads to the GMB heuristic requiring additional time periods to reach full vaccination of the population after the MC heuristic has already stopped vaccination. It is unclear why exactly the GMB heuristic behaves this way, but it can be observed that the difference between how many vaccines are sent to a zone for the two strategies is usually less than 2000, i.e., the size of 1 lot of vaccines. Thus, one reason for the behavior of the GMB might have been because the base ten proportioning scheme was used, and the GMB heuristic was allocating doses in units of 10000, 1000, 100, 10, and 1. When allocating 1000 or fewer doses at a time, the first doses would still require sending a whole additional lot to that zone. Therefore, if the benefit of sending a partial lot of vaccines does not outweigh the costs, the GMB heuristic will not send the lot.

However, this idea has a problem. Up to that point, the GMB heuristic would have been allocating partial lots at a time, and it always decided that such an action was worth the cost. Why does it suddenly change so close to the administrative capacity limit? In fact, up until the point that allocating a portion would cause a region to exceed its administrative capacity, the algorithm has no way of knowing what the administrative capacity of a region is. Perhaps once the finite number of 1000 dose portions are exhausted, when allocating 100 dose portions, the benefits from such a small fraction of a lot are outweighed by the costs of sending a full additional lot. The true cause of this behavior will require additional study to identify. An experiment to test the previous hypothesis would be to observe whether the GMB will allocate up to the administrative capacity if there were to be no downside to sending additional vaccines, such as by setting the weight of the costs in the objective function to zero.

Table 5.2 shows the run time of each solution method for instances of varying sizes, from which we can make several observations. First, it is clear that the MC heuristic is, by far, the quickest solution method, outperforming all other methods for every tested instance. It can be seen that while, in general, the run time increases with problem size, this is not always the case for the MC heuristic, such as with instances 18 and 20 with run times of 0:02:35 and 0:00:45, respectively. Similarly, we can see that the run time appears to increase and decrease randomly with supply, as seen in problems 23, 37, 38, 39, and 40. This behavior can most likely be attributed to the looping behavior of the algorithm, where sometimes the algorithm would bounce between several similar solutions before the algorithm detected that it was in a loop, while other times it quickly converges onto a single solution. Meanwhile, by examining the GA, its run time generally increases with problem size. Besides the stopping condition, the primary factors impacting the GA run time are the time it takes to evaluate a chromosome and the number of chromosomes to evaluate in each generation. For the GMB heuristic, the primary factors that would impact run time

would be the number of portions, the number of potential allocations it has to check per portion (i.e., $|I| \times |G|$), and how long it takes to evaluate a potential allocation. It can be seen that thanks to the proportioning scheme used, the run time does not tend to increase much alongside increasing supply, as the number of portions to consider increases slowly with the overall supply. For both the GA and the GMB, in order of impact on the run time, it seems to be the number of 1) groups, 2) zones, 3) vaccine types, 4) time periods, and 5) DCs, with the number of DCs having very little impact. It is worth noting that we are increasing the considered time periods from 10 to 20 to 30 in these tests, which makes the differences from time seem larger compared to the other metrics.

From Table 5.6 we can get an estimate of the optimality gap for the GA+GMB solution method. The gap between the LB identified by GUROBI and the solution found using the GA+GMB solution method ranges from 0-14.25%, though most problems are less than 5%. These values represent the absolute maximum optimality gap that could exist between the GA+GMB solution and the true optimal. However, there is reason to suspect that the true optimality gap is likely lower. If we look at the gap between the UB identified by GUROBI and the solution found using the GA+GMB solution method, it is consistently less than 2%. This represents the absolute minimum the optimality gap could be. From this and observing the value of the UB and LB over time while running GUROBI, where the UB does not change very much and most of the run time is spent slowly improving the LB, it becomes clear that the LB obtained from GUROBI is usually worse than the UB. For example, in problem 2, there is a 13.28% gap between the GA+GMB solution and the LB, but only a 0.14% difference between the GA+GMB solution and the UB. Thus, at least for these small problems, it would suggest that the approximate optimality gap when using the GA+GMB solution method is closer to 2% than 14%. These problems are small, and its almost certainly true that the optimality gap would grow with the problem size. However, it should not be much larger, so a 5-10% or less optimality

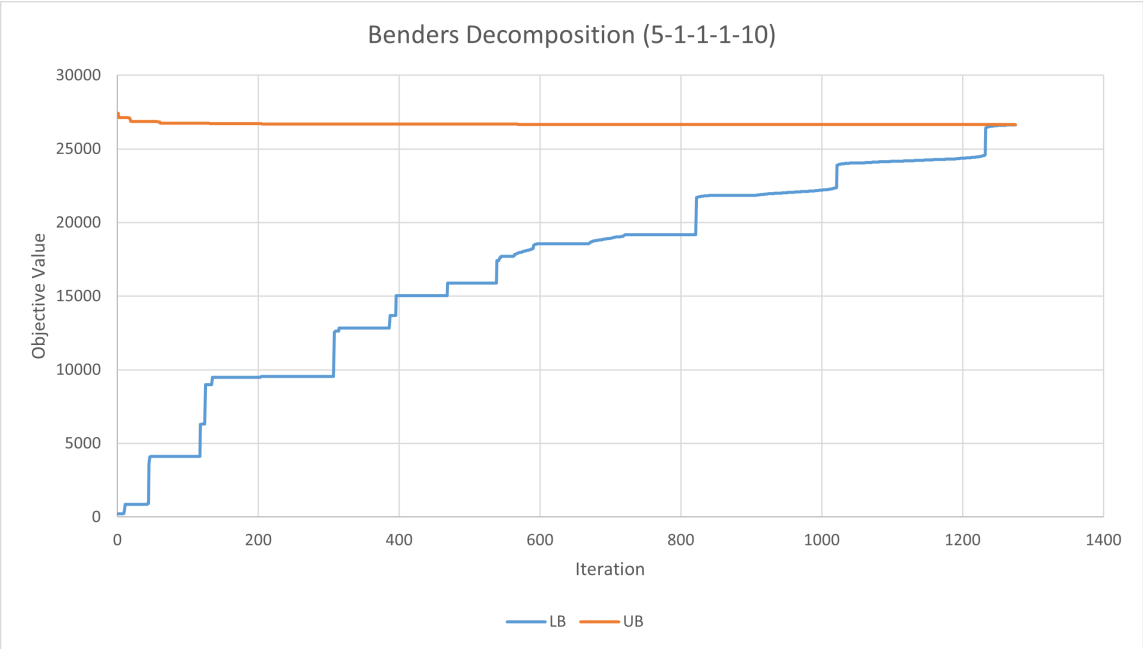
gap would seem to be a fair estimate for the GA+GMB solution method.

In regards to BD (Table 5.7), it was only able to be run on problems with just a few potential lots of vaccines, and with only a few zones and time periods. Even for these small problems, BD takes a prolonged period to solve compared to the other solution methods. However, in terms of the objective value, it always either matches or outperforms the other solution methods. These problems are sufficiently small that a commercial solver can solve them, which of course defeats the purpose of using the BD solution method, but does allow us to verify that Benders does indeed produce the optimal solution. It can be observed that with increasing zones and time periods, the run time and iterations required to arrive at the optimal solution grow rapidly. Crucially, however, it does not appear that the number of iterations is the same as every possible combination of allocating the supply of vaccines. For example, if every possible solution were explored for 5-1-1-1-10 with 1 lot of vaccine available each time period, with 6 potential ways to allocate the lot per time period and 10 time periods to consider, that would be 60,466,176 possible ways to allocate the lots, but only 1,233 sets of cuts were required. This means that logic-based cuts can eliminate multiple possible solutions at a time. If these cuts could be generated quickly, they would prove incredibly useful. However, it was found that as more cuts are added, building and solving the Benders master problem takes more and more time, until multiple minutes are required per iteration. As currently implemented, the model is built from scratch and solved using the GUROBI solver each time. If the build and solve time of the Benders master problem could be improved, then potentially the speed of the cut generation could outweigh the number of cuts required and allow Benders to be used for larger problems. However, it will still only ever be useful in single-risk group problems.

Figures 5.3 and 5.4 depict the convergence of the UB and LB of the BD method for two of the test problems. Figure 5.3 only had a single lot per time period to allocate and it can be observed that steady progress was made in improving the LB

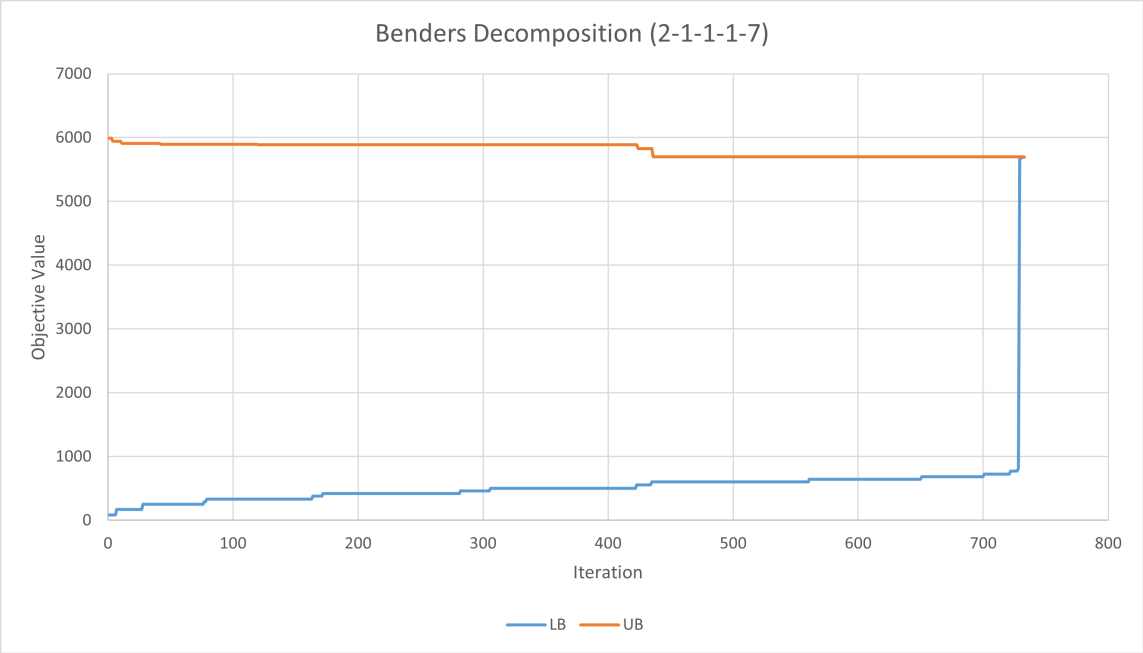
overtime. Figure 5.4 meanwhile had two lots of vaccine to allocate, and it can be seen that convergence was slow up until the last few iterations where a sudden sharp jump in the LB brought the two together.

Figure 5.3: Benders Decomposition UB and LB (5-1-1-1-10)



These results also confirmed one of the factors that makes combining the vaccine allocation and distribution models more difficult to solve in comparison to the regular vaccine allocation problem. Because the problem is now multi-objective, there are potentially multiple different strategies that arrive at the same objective value. This was best highlighted by two test runs using the PR and GMB heuristic on a small problem. Table 5.9 shows the combined objective value, the two components of said objective value; total new exposures and costs, as well as the total vaccines allocated. The objective functions of these two methods are very similar, differing by only 1,636. However, they arrive at this value from different directions. The PR allocation uses a little over a million vaccines, and as such prevents many exposures at

Figure 5.4: Benders Decomposition UB and LB (2-1-1-1-7)



great monetary cost. The GMB heuristic meanwhile allocates only 20,000 vaccines, leading to many more exposures, but making up for it in terms of costs. Thus, two completely different strategies may yield very similar results.

Table 5.9: Method Comparison: Similar Results

Method	Objective	Total NE	Total Costs	Total Vaccines Allocated
Pro-Rata	29,303	7,682	$ 21,621 000	1,029,186
GMB	30,939	30,477	$ 462,000	20,000
Difference	1,636	22,794	-$ 21,159 000	-1,009,186

5.3 Method Comparison

This section will review the performance of each solution method, including those that ended up not working as intended, and for such methods, review why they did

not work, as well as whether there are potential avenues to explore them further.

5.3.1 Lagrangian Relaxation Performance

After exploring Lagrangian relaxation as a solution method, it was deemed to be unsuitable for the task. The problem is two-fold. First, when constraint (3.32) is relaxed, the integrity of the model is severely damaged, as the resulting models' objective functions all compete with each other. SP1 seeks to minimize the costs of distributing vaccines, which naturally means it wants to allocate as few vaccines as possible. For SP2, each zone will seek to minimize its own epidemic impact by allocating as many vaccines as possible, even if those vaccines do not exist or another zone needs them, as the sub-problems have no connection to the supply chain or each other anymore. It was conjectured that λ should work to correct these differences over each iteration of the algorithm. However, experimentation has shown that this does not occur, and instead, the algorithm will never rectify the competing interests between the sub-problems.

The second issue is that with this method SP2 is still difficult to solve in its own right. SP2 is still the non-convex vaccine allocation problem, and in fact in order to run the algorithm, the sub-problem can often only be solved to obtain an upper bound on its optimal solution, which severely impacts the quality of the final Lagrangian bound. However, if the problem is simplified to a single risk group (i.e., a homogeneous model), then SP2 can be solved through simulation and enumeration. While this issue could be rectified if some efficient method could be identified to solve SP2 to optimality, such a method could likely just as easily be applied to the original problem. This is the crux of the issue with Lagrangian relaxation. It does little to actually address the issues that make the problem difficult to solve.

5.3.2 McCormick Envelope Performance

As previously stated, the performance of the McCormick Envelope method would be dependent on the quality of the upper and lower bounds that could be obtained on decision variables S_{igt} and I_{igt}. Unfortunately, no good bounds on these variables could be determined. An upper bound on the infected population could be obtained by running the model without allocating any vaccines, as there will never be any more infected than this, as allocating vaccines can only serve to lower the number of infected, not increase it. However it is not a very good bound and, unfortunately, there does not seem to be a tighter upper bound, or at least one that can be obtained easily. This is because if no vaccines are allocated to a region, which is a very plausible outcome, then only the upper bound of the no-vaccination strategy will hold. The lower bound is similarly very loose, as the only guaranteed infected are those who start out infected and those who will always become infected in the first period before vaccination begins. A tighter bound could be found if it could be proven that exclusively vaccinating one group over another will always lead to lower overall cases for the first group, as then a lower bound on the infected population could be said to be the number of infected in a time period if up to that point, all resources had gone towards vaccinating that specific group each time period. However, there is the possibility that this assumption would not hold in all scenarios. For example, consider a scenario with just two groups, one with more contacts per week than the other. If we do not vaccinate the group with the larger number of weekly contacts, the disease spreads out of control through that group. Even though a portion of the other group has been vaccinated, there are now so many additional infected from the first group that the protection is less than if we had vaccinated the first group. In this way, a group can have more infected contacts under a scenario where they were vaccinated than if some other group was vaccinated first.

A similar problem exists for the susceptible populations, as there is no good

upper or lower bound on the susceptible population. Consider the scenario of vaccinating no one. There might still be some susceptible population left over at the end, those who got lucky and never caught the disease. Then consider allocating only one additional vaccine somewhere. The final susceptible population of that group is decreased by one, but all groups in the zone will collectively receive a small increase in protection due to the herd effect. Thus by vaccinating one group, the final susceptible population in another group grows. If the herd effect is strong enough, i.e., each vaccine prevents more than one additional infection, the final susceptible population for the group it was allocated to might also increase. This means you can end up with a larger susceptible population than when not vaccinating, thus no upper bound can be obtained on the susceptible population. Similarly, the lower bound on the susceptible population is not very good. By sending no vaccines, we can obtain the susceptible population of a group in each period that would not have caught the disease yet, which is the lowest the susceptible population can go without vaccinating the group directly. Because vaccinating the group will lower their susceptible population, the lowest the susceptible population of a given group can be in a period is the susceptible population when letting the disease run its course minus the vaccinated population of the group if all vaccines had been allocated in all previous periods. This is a very loose bound that would only get worse with each additional period considered in the model.

These poor bounds meant that the McCormick envelope would not produce a very good approximation of the original problem.

5.3.3 Mountain Climbing Heuristic Performance

The MC heuristic solution method initially seemed to be very promising as a solution method. As can be seen from the results on some of the small to medium-sized problems, it frequently outperforms the other algorithms in both objective value and solution time. In fact, it is the fastest-solving method in all scenarios tested.

However, what is not obvious from the raw results is that for the majority of the MC tests, the solution did not converge perfectly. Instead, it would alternate between several solutions, of which the best performer was taken as the final solution. For those problems where it was the best performing method, it was usually only barely better than methods such as the GMB heuristic, with less than a 5% difference at most, though it was able to arrive at solutions significantly faster than the GMB heuristic (e.g., Problem 11: MC 0:00:14 GMB 0:03:26).

However, on larger problems the method would seem to begin to break down, converging on local optimum far from the other solutions. What is confusing is that there was no obvious reason or pattern for this change. For example, as can be seen in Table 5.5, for problems 11 (5-1-3-1-20) and 23 (10-1-3-1-30) the method worked well, improving on the GMB heuristics by 4.22% and 3.74%, respectively. Meanwhile, for problems 8 (5-1-1-1-20) and 24 (10-1-5-1-30) it failed, producing objective values that were 16.32% and 5.74% worse compared to the GMB heuristic, respectively. Therefore, although its speed and performance on some problems were very impressive, the MC heuristic proved to be unreliable as a solution method, especially for medium to large problems. However, if the reason for this inconsistency could be identified and corrected, it might prove to be a very promising solution method. Perhaps if the algorithm could be started with a near-optimal solution, such as from the GMB heuristic, it would avoid this issue of falling into a poor solution, though this would come at the expense of the method's primary benefit, its run time. Alternatively, it might prove to be useful as a method to warm start the GA. Even though the solutions it produces are potentially worse compared to the GMB heuristic, the speed at which they can be derived combined with the improvement from the GA could make up for the drop in quality of the warm start solution.

Figures 5.5, 5.6, and 5.7 show the course of the MC heuristic over its run time on three different problems. The first, depicted in Figure 5.5, is on a small problem that

reached a good solution very quickly, without getting stuck fluctuating between solutions. Figure 5.6, on the other hand, shows the algorithm became stuck fluctuating between two points, 2,495,346.299 and 2,485,904.886. While it may appear that it entered this state after just 6 iterations, there is actually always a small difference between each point up until iterations 15 and 17, which produced completely identical solutions, at which point, detecting a loop, the algorithm terminated. Additionally, it can clearly be seen that the best-performing solution was actually from the second iteration, and the algorithm moved away from that solution with subsequent iterations. Figure 5.7 displays a similar pattern, but instead of fluctuating between just two solutions, it is fluctuating between several solutions. However, unlike the previous problem, none of these points are actually exactly the same, so the algorithm did not stop on its own. Instead, it was manually stopped after 100 iterations. If it was allowed to run longer, it might have eventually produced a solution that was identical to the previous one and entered a loop. After all, as was observed with the previous problem, it can take a few iterations before settling into a loop, and perhaps larger problems take more time to settle into a pattern.

Figure 5.5: MC Solution Over Time (5-1-3-1-20)

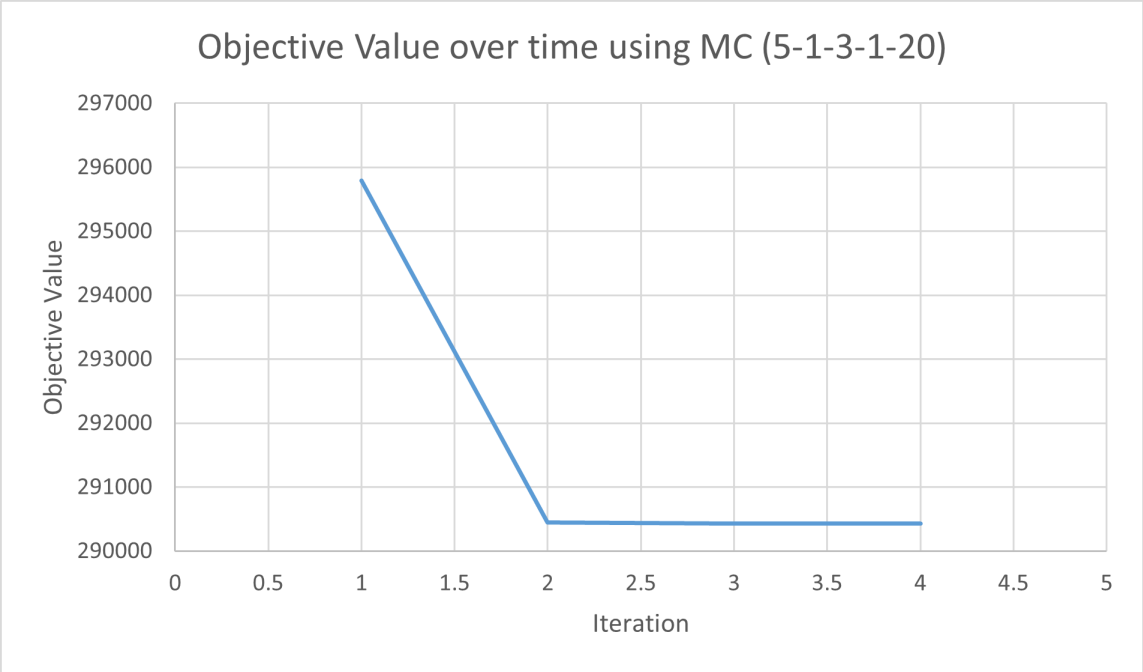


Figure 5.6: MC Solution Over Time (10-1-5-1-20)

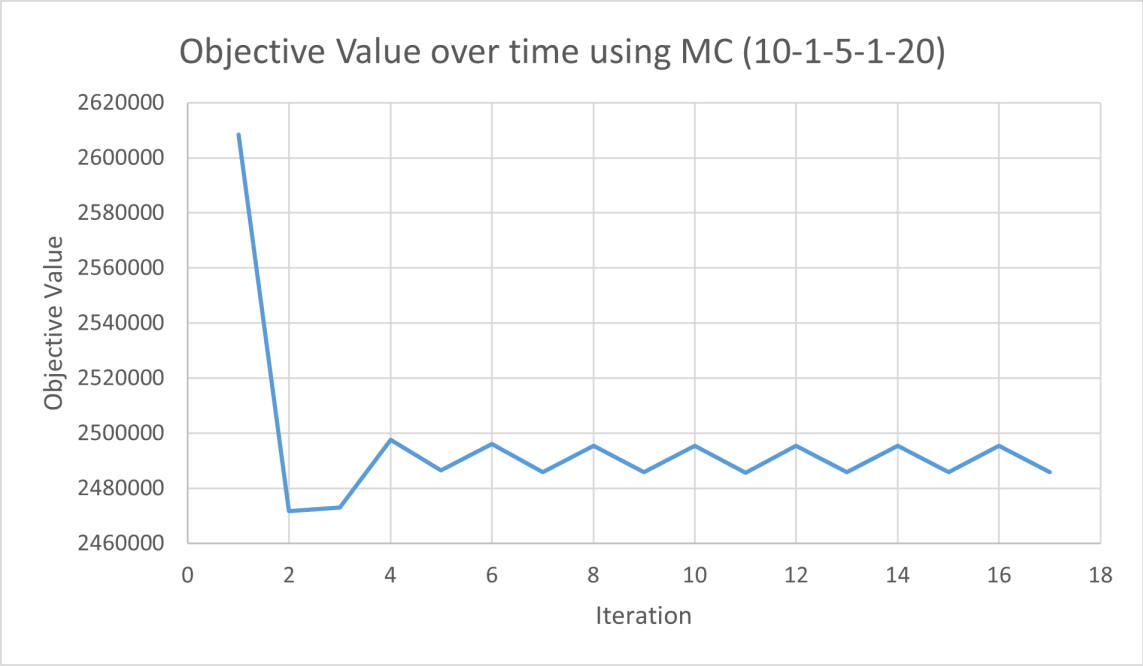


Figure 5.7: MC Solution Over Time (20-1-5-1-20)

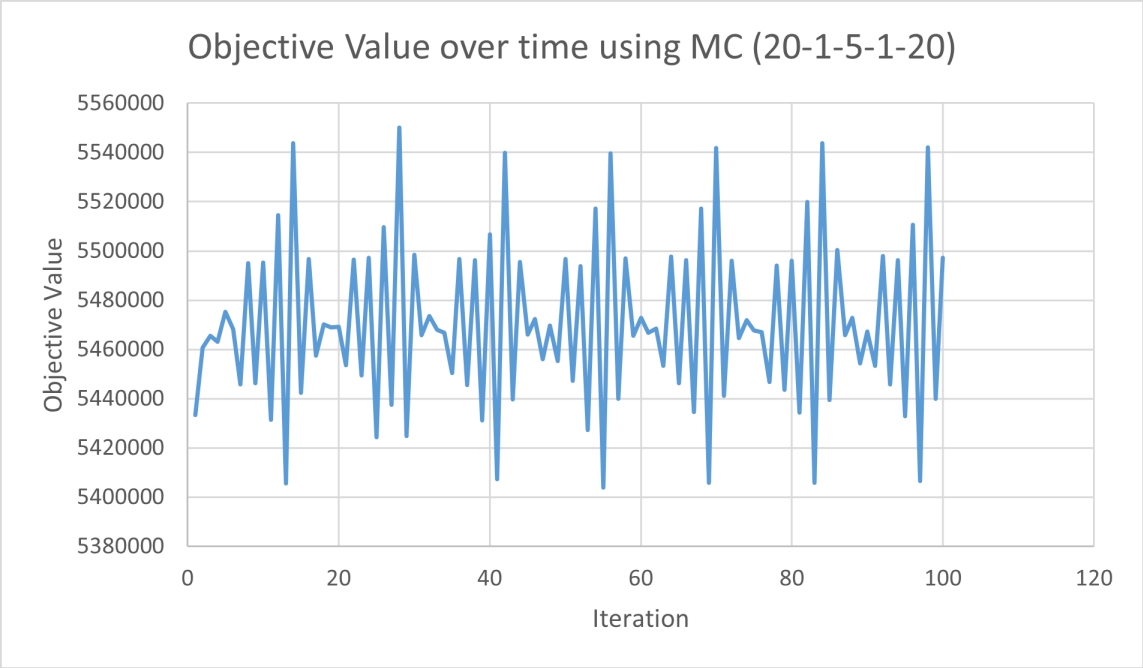
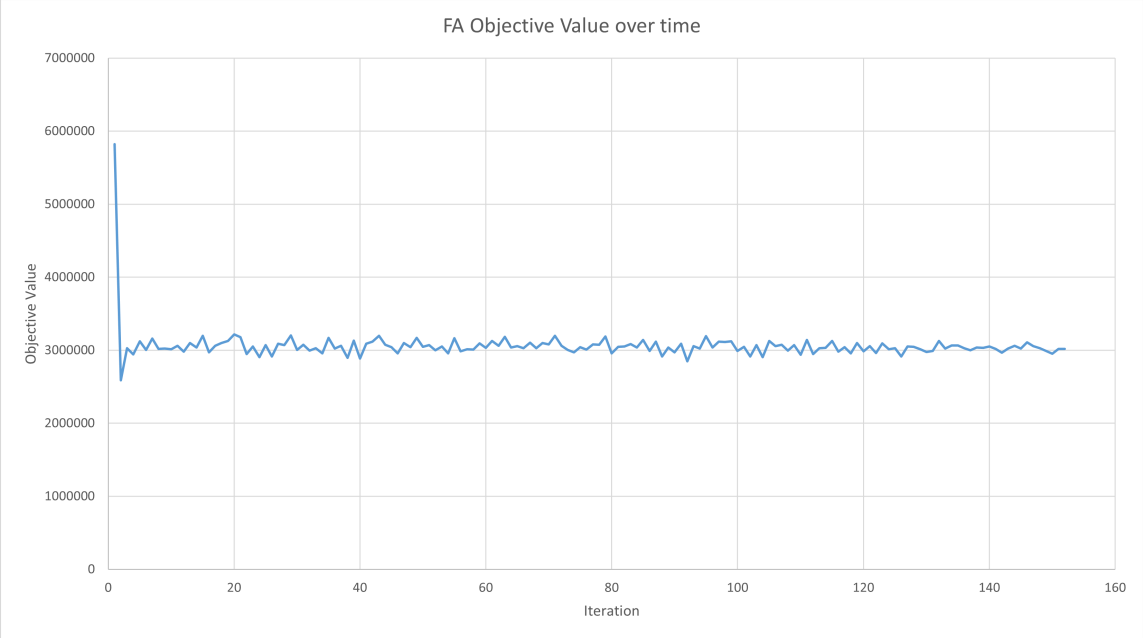


Figure 5.8: Fixed Allocation Algorithm Solution Over Time



5.3.4 Fixed Allocation Linear Approximation Performance

It was found that solutions generated by the fixed allocation linear approximation solution method were not very good, and the algorithm would often get stuck alternating between a few different solutions far from the optimum, never able to outperform the corrected Pro-rata strategy. On top of that, it is not a very fast solution method and requires keeping track of whether the algorithm has got itself stuck in a loop to determine when to terminate. Another problem that can be seen by observing the objective value for each iteration is that the solutions do not appear to get better over time. As can be seen in Figure 5.8, for the test scenario of 10-1-5-1-20 with 200 lots of vaccines, the best-performing solution was actually the second solution the algorithm generated. It never managed to return to that value and does not appear to be trending toward it.

In fact, at the point of termination, many better solutions had previously been

identified. This pattern was consistent for all tested scenarios. In theory, the best-performing of these solutions could have been selected and treated as the final solution. However, such solutions are basically equivalent to choosing a solution at random. There is no evidence to suggest that the best-performing solutions generated over the course of the algorithm's run-time are anywhere near optimal.

However, the method of deriving the linear formula for the new exposures from the fixed allocation was not very sophisticated. It simply used the slope between the new exposures and the susceptible population (see Equation (4.42)). In comparison to other methods, less time was spent on the fixed allocation method when other solution methods already showed promising results. If a better method of linearizing from the results of the fixed allocation strategy could be found, the solution method as a whole might be improved.

5.3.5 Piecewise-Linear Approximation Performance

The piecewise-linear approximation solution method was found to only complicate the problem, rather than help solve it. In order to maintain a reasonably low approximation error, depending on the initial susceptible population, thousands or even tens of thousands of breakpoints and decision variables are required, for each zone, group, and time period. The total number of breakpoints and SOS2 decision variables required can be calculated using equations (5.2) and (5.3), respectively. Introducing tens of thousands of SOS2 binary decision variables obviously makes the problem difficult to solve, likely more so than the original formulation with the additional disadvantage of now only being an approximation of the original. Whether these issues can be remedied is unclear, but if not, then piecewise-linear approximation is not a good approach to trying to solve the vaccine allocation problem.

$$\#\text{ Break Points} = \sum_{i \in I} \sum_{g \in G} \frac{S^0_{ig}}{2\sqrt{\epsilon}} \tag{5.2}$$

$$\#\ \text{SOS2 Decision Variables} = |T| * |V| * \sum_{i \in I} \sum_{g \in G} \frac{S_{ig}^0}{2\sqrt{\epsilon}} \tag{5.3}$$

5.3.6 GMB Heuristic Performance

Being a heuristic method, the GMB solution method cannot guarantee an optimal solution, only a good feasible solution, even when taking the most conservative approach possible by allocating one vaccine at a time. Experiments done by varying the vaccine supply showed that the amount of vaccines available will impact where vaccines will be most effective. For instance, if there are only 100 doses, it might be best to vaccinate group X but if there were 1000 doses it might be better to vaccinate group Y or to split them between several groups. Because the algorithm starts from the scenario of no vaccines being allocated and allocates starting from $t = 0$ forward, it considers the allocation of each portion of the vaccines in isolation, and thus only allocates vaccines where the most immediate benefit is found, without consideration for further vaccines or whether its previous allocations might no longer be the best. Thus, by only ever considering portions of the vaccine supply at a time and not the entire supply over the entire time period, the algorithm makes imperfect decisions.

A shortcoming of this approach is that it cannot make use of the storage capacity of the DCs. Vaccines must be used in the period they are made available. This has no impact if the vaccine supply is steady and there is always enough capacity to distribute and administer all available vaccines. However, in the event that vaccine supply fluctuates, i.e., there are periods of high and low supply, and during the high points the supply outstrips capacity, those vaccines go unutilized if not stored in a DC.

The size of the problem determines how long simulating each allocation takes, and the number of zones and risk groups determines how many simulations are required per allocation. The supply of vaccines and in what size portions they are allocated determines how many allocations will need to be performed. Thus the speed and

performance of this method are directly proportional to the supply of vaccines, the portioning scheme used, the number of zones and risk groups, and the size of the problem. The run time can thus be calculated using equation (5.4), making it an $O(n^3)$ algorithm.

$$\text{Run Time} = ((\text{SEIR Solve Time} * |G| * |I|) + (\text{Transshipment Solve Time} * |I|)) * \#\text{ portions} \quad (5.4)$$

The run time; while deterministic and faster than the GA for all tested scenarios, is ultimately dependent on the number of available vaccines and the proportioning scheme used. While using an appropriate proportioning scheme can significantly cut down on the number of evaluations the algorithm will need to perform, for very large quantities of vaccine, there is potential for the GA to be faster in terms of solve time, as its solve time is independent of the vaccine supply.

Additionally, as previously mentioned, the solution quality is dependent on the proportioning scheme. In general, a more granular search (i.e., more, smaller, portions) will lead to a better solution compared to a less granular one at the expense of time. However, where vaccines will be most effective depends on how many vaccines are available. The concept of the herd effect and the dose-optimal vaccine fraction [20] is relevant here, as the benefit of multiple vaccines is more than the sum of its parts. By allocating one vaccine at a time, the compounding effect of multiple vaccines might be lost, and thus if sending 100 vaccines to group X would lead to a better solution in the end but each individual vaccine is more immediately effective elsewhere, the one-by-one allocation will not capture this effect. The same problem exists for the larger allocation sizes as well. Consider if when allocating a 100-dose portion, the most benefit would be found at 50 doses or 500 doses. Thus, it seems likely that there is no optimal proportioning scheme to allocate vaccines using this method, or if there is one, finding it might be just as difficult as solving the original problem. There is also the problem that when allocating portions that are smaller than the size of a lot of vaccines, allocating just a few vaccines will still necessitate

Table 5.10: Proportioning Scheme Comparison

Proportioning Scheme	Objective Value	Solve Time
Bulk Starting from 10,000	2,364,281	0:38:43
Bulk Starting from 100	2,500,442	7:14:46
Fewest Portions	2,357,433	0:55:19

sending an entire additional lot to the zone. Sending a whole lot and only getting a fraction of the benefit might not be worth the incurred costs, especially for larger values of β. This causes the algorithm to not allocate any more vaccines to that region, whereas if it had continued to allocate to the region it would no longer incur additional costs as the lot was already going there anyway. Table 5.10 displays the objective value and solve time for three different proportioning schemes used on the same problem.

The more granular proportioning scheme starting from allocating 100 dose portions actually yielded the worst results of the three tests. Meanwhile starting from allocating 10,000 vaccines at a time ended up yielding a better final objective value, while the fewest portions did better still. These results demonstrate that while in general, a more granular proportioning scheme is better, this is not always the case. One idea worth exploring with the GMB heuristic method is testing different portion sizes for each allocation, measuring the benefit per vaccine, and choosing the allocation that provides the most benefit per vaccine used. This would help address the issue of not allocating in the correct portion size. However, it would likely slow down the heuristic considerably, multiplying the number of evaluations that need to be performed per allocation by the number of potential portion sizes. Given that it already performs well, this trade-off in run time might not be worth it, especially if instead the GA can be used to improve its performance.

Through experimentation, it was found that, on its own, the GMB heuristic is the best-performing solution method for medium to large problems. For small problems, it is only surpassed by the MC heuristic, where it performs comparably in terms of

solution quality, but slower in terms of run time. While additional groups, regions, time periods, and an increased vaccine supply will lead to more opportunities for miss allocating a portion of vaccines, harming the solution quality, the GMB heuristic has been shown to more reliably produce near-optimal solutions, even for these larger problems, compared to the other solution methods.

5.3.7 GA Performance

As the GA method has no known guarantee to determine if it has reached the optimal solution or not, the quality of the objective value and time to solution is highly dependent on what the specified stopping condition is for the algorithm. The stopping conditions used for these experiments were imposing a limit on the number of generations or observing less than L (1) improvement over the previous K (100) generations. How long it takes to reach the termination condition depends primarily on the size of the problem ($|I|, |J|, |G|, |V|, |T|$), as this impacts both the evaluation time of each chromosome and the number of genes that need to be considered. When running without a warm start from the GMB heuristic, the GA takes a prolonged period of time just to arrive at the same objective value as the heuristic on its own, with the experiments confirming that the GA never manages to outperform the GMB heuristic within the heuristic's run time. Conversely, thus far there has not been a scenario where the GA was unable to improve on the solution generated by the heuristic. The improvement from the heuristic solution is usually small relative to the objective value of the heuristic, often representing less than a 1.5% improvement. However, this small improvement is indicative that the heuristic method provides a fairly good upper bound on the optimal solution, as a near-optimal solution is harder for the GA to improve on. This combined approach represents the best solution method identified in terms of objective value, and as demonstrated in Table 5.7 can sometimes produce the optimal solution. However, it is still only a heuristic, with no guarantee or even expectation of optimality, and for larger problems, there

is no method of verifying optimality.

Figure 5.9 depicts the fitness score (i.e., the objective value) of the top performing chromosome of each generation when using the GA alone, with Figure 5.10 displaying the improvement from the previous generation. Figures 5.11 and 5.12 are the same graphs for the case where the GA was given a warm start by the GMB heuristic. Without the warm start, the algorithm was seeing gradual improvement with each subsequent generation. By the time the stopping condition was reached (1000 generations) there was still room for improvement to be made and according to Figure 5.10 the rate of improvement even appeared to be accelerating. With a warm start, improvement can be seen to be more sporadic and gradual, with periods of very small improvement interspersed by large jumps in the objective value performance. By the time of termination, the algorithm had largely stopped making these occasional large improvements, though there is no way of knowing how much potential for improvement there still was if the algorithm was allowed to run longer. Based on the results from Table 5.6, the optimality gap is likely to be 5-10% at most, likely smaller, especially for problems that are not exceedingly large.

Figure 5.9: GA Solution Over Time

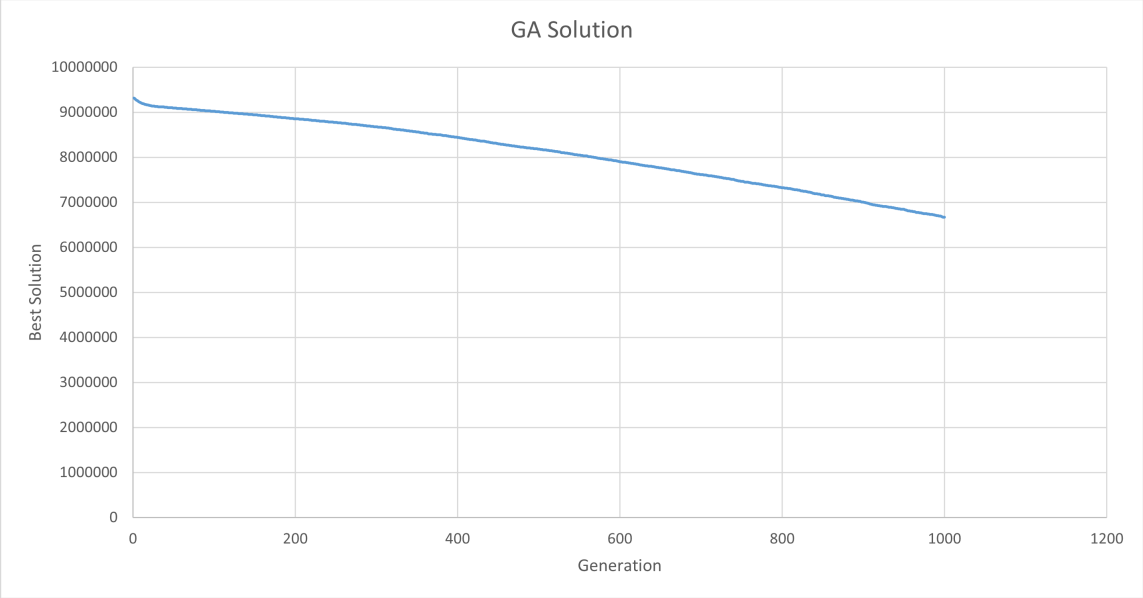


Figure 5.10: GA Improvement Over Time

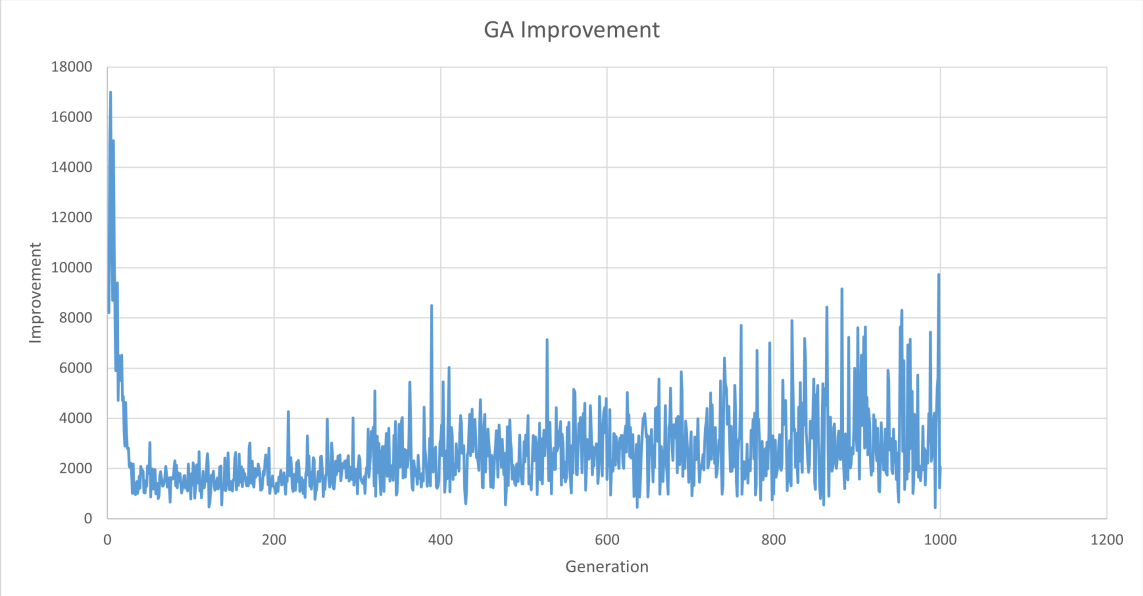


Figure 5.11: GA with Warm Start Solution Over Time

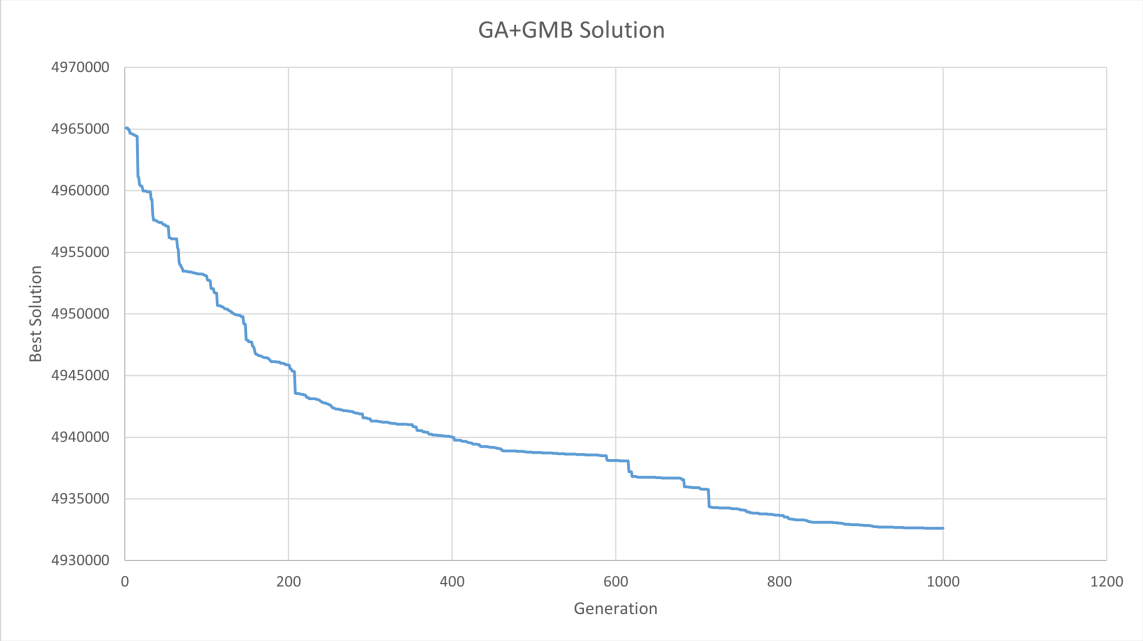
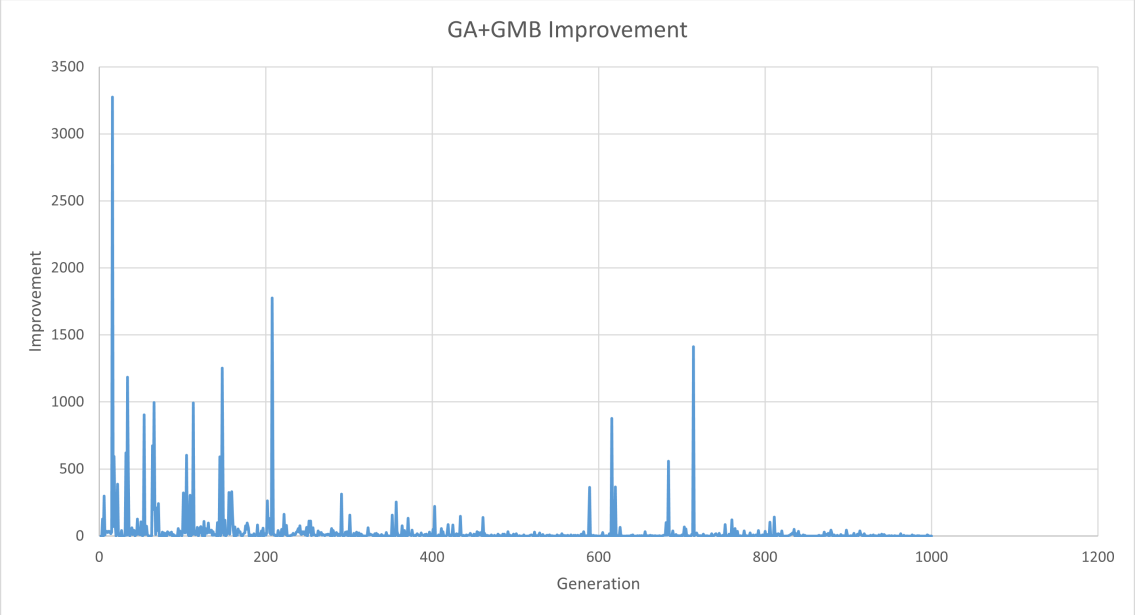


Figure 5.12: GA with Warm Start Improvement Over Time



Even without the use of the GMB heuristic, the GA has demonstrated that it can relatively quickly identify good solutions to the vaccine allocation and distribution problem, though not of the same quality as either heuristic method. Given sufficient generations, the GA will invariably arrive at a near-optimal solution. When combined with the GMB heuristic, it has been able to generate the best solutions of the solution methods explored in this thesis thus far. Unfortunately, as the algorithm approaches the optimal, improvement slows down drastically, as increasingly specific crossover and mutation events are required to occur to improve further. However, the fact that the GA can no longer improve on a solution, while not a guarantee of optimality, does at least suggest the solution is near-optimal. As the GA makes minimal improvements to the GMB heuristic, this also supports that the GMB heuristic on its own has likely provided a near-optimal solution to the vaccine allocation problem. It would be worthwhile exploring the benefits of using alternate methods of generating the warm starts for the GA. For instance, the MC heuristic may not provide quite as good a starting solution as the GMB heuristic, but it can find it in a fraction of the time. The hope would be that the GA could improve the solution from the MC heuristic to the point it is comparable to the GMB in less time than it would take to run the GMB heuristic itself.

One potential method of improving the performance of the GA would be to further tailor the mutations towards producing genes that tend to show up in optimal solutions. For example, the optimal strategy often contains periods where the focus is on a few select groups, administering as many vaccines as possible to them which usually entails administering up to the administrative capacity of a zone. However, it is difficult to arrive at the administrative capacity level through random chance alone. Making some portion of the mutations set the value of a non-zero gene to the administrative capacity and all other genes in the region to 0 would likely help to increase the likelihood of such genes being produced in the algorithm's run time. As currently formulated, the GA also has the potential to waste small portions of

the final lot of vaccines sent to a zone each period. This is because it would require an allocation that is a multiple of the lot sizes, which the random mutations of the algorithm do not account for, and thus not all the vaccines that get sent to the region might get used. This obviously hurts the overall solution as vaccines are being distributed to a region but not being utilized. This problem could be corrected by rounding up the allocations in a region to the nearest multiple of the lot sizes, though there would have to be some system in place to decide which groups received these additional vaccines. However, this is likely a relatively minor issue, as given sufficient generations and mutations, the algorithm will solve this issue on its own, as mutations that shrink the gap between the number of vaccines distributed and the number of vaccines allocated will be found to be beneficial and thus propagate through the population. Other similar mutation rules could be developed to further improve the chance of beneficial genes occurring. However, caution should be taken to not introduce too much bias into the selection process that would tend to lead it toward these solutions. This would defeat the original purpose for using the GA to explore the solution space beyond what the heuristic could and consider alternative strategies.

5.3.8 Benders Decomposition Performance

As previously mentioned, the benefit of the BD approach is the guarantee of optimality it provides. Each logic-based cut generated is an optimality cut that will reduce the solution space and it can be proven that this method will always eventually return the optimal solution. However, the logic-based cuts are generally not very good cuts, potentially only excluding a single solution from the solution space. In the worst-case scenario, it might need to exhaust nearly every possible solution (i.e., every possible configuration of dividing X lots of each type of vaccine $|V|$ amongst $|I|$ zones and $|G|$ groups, for every time period $|T|$) taking a prohibitively long time to solve to optimality. In addition to the poor quality of the current cuts, it was

found that as the number of cuts grows, the Master problem's build and solve time slows down significantly, often taking several minutes to solve a single iteration of the Master problem. Therefore, the run time of the algorithm for all but the simplest of problems is intractable.

The core problem with BD as a solution method though is that, similar to Lagrangian relaxation, it does little to solve the issues that arise with trying to solve the sub-problem, such as the non-convex constraint (4.74). If there are multiple risk groups, the sub-problem is not made much easier to solve through the use of BD. Some secondary solving method is required to solve the sub-problem, which if it works, could likely just be applied to the original problem. However, because the master problem decides how many lots a zone receives and not using a vaccine when it has been delivered to a zone is never beneficial, the optimal policy is to administer all delivered vaccines up to the point they exceed the administrative capacity of the region. Thus, when there is only one group in each zone to receive those vaccines, the optimal policy is to administer all vaccines to that one group and the outcome of administering the entire supply of vaccines to said group can be found efficiently using simulation. Therefore for problems with only a single risk group but multiple independent zones, the sub-problems can be solved using simulation, allowing BD to more easily solve such problems. This means that while this BD solution method can, in theory, be used on the heterogeneous vaccine allocation problem, in practice it is only really beneficial when used to solve the homogeneous vaccine allocation problem as otherwise another solution method is required to solve the sub-problem to the required guaranteed optimal.

If a method of generating improved cuts could be identified and the build and solve time of the master problem could be reduced, the BD approach shows promise as a method of solving the vaccine allocation problem for the case of a single risk group with multiple zones. However, for problems with multiple risk groups, BD does not address the issue of the non-convex allocation problem, and so does not

provide much benefit for using it.

5.4 Method Comparison Summary

To summarize, the two best solution methods are the MC and GMB heuristics, with up to 85% improvement over the NI policy and 46% improvement over the corrected Pro-rata strategy. The MC heuristic is the fastest solution method, and it performs comparably to the GMB heuristic in terms of solution quality for small problems ($<$ 4% difference) but gets outperformed for larger problems ($<$ 16% difference). By using the GMB heuristic to warm-start the GA, the solution can be improved further, though the benefit is often minimal for the additional run time ($<$ 4% improvement from the heuristic alone). While neither the MC nor GMB heuristic can guarantee optimality, the fact that the GA cannot improve on the objective value supports the idea that their solutions are at least near-optimal. With experiments on smaller problems for which GUROBI can be used to obtain a LB suggesting that the optimality gap is likely 5-10% or less. BD can be used to solve problems with a single risk group to the guaranteed optimality, but with the current method of generating cuts, it is infeasible to do so for all but a few select problems.

Chapter 6

Case Study

This section applied the most promising solution methods explored in this thesis to a case study of vaccine allocation and distribution during the COVID-19 pandemic in the province of Ontario, Canada. Ontario was chosen for two reasons. First, from the literature review, there is yet to be a study that has used a Canadian COVID-19 case study. Second, Ontario has a wide array of publicly available data on the COVID-19 pandemic that makes it easier to develop a case study for it [52]. Pioneering vaccines were first authorized for public use in Canada in December of 2020, with widespread vaccination beginning shortly after. The case study will focus on this early period, when few were vaccinated, and vaccines were in limited supply. Deciding how to make the best use of these limited vaccines was a crucial decision during this time frame.

6.1 Case Study Data

Ontario divides its health care service into 34 Public Health Units (PHU) and publishes various public health data sets broken down by PHU. PHUs cover specific non-overlapping geographic areas. This feature makes the PHUs very convenient for use as the population zones in the case study, as in this way, the entire provincial population can be accounted for, not just major urban population centers. Figure 6.1 displays a map of the PHUs in Ontario colour-coded based on population.

Data is available on the population of each PHU broken down into 9 age groups: 5-11, 12-17, 18-29, 30-39, 40-49, 50-59, 60-69, 70-79, and 80+. For the case study,

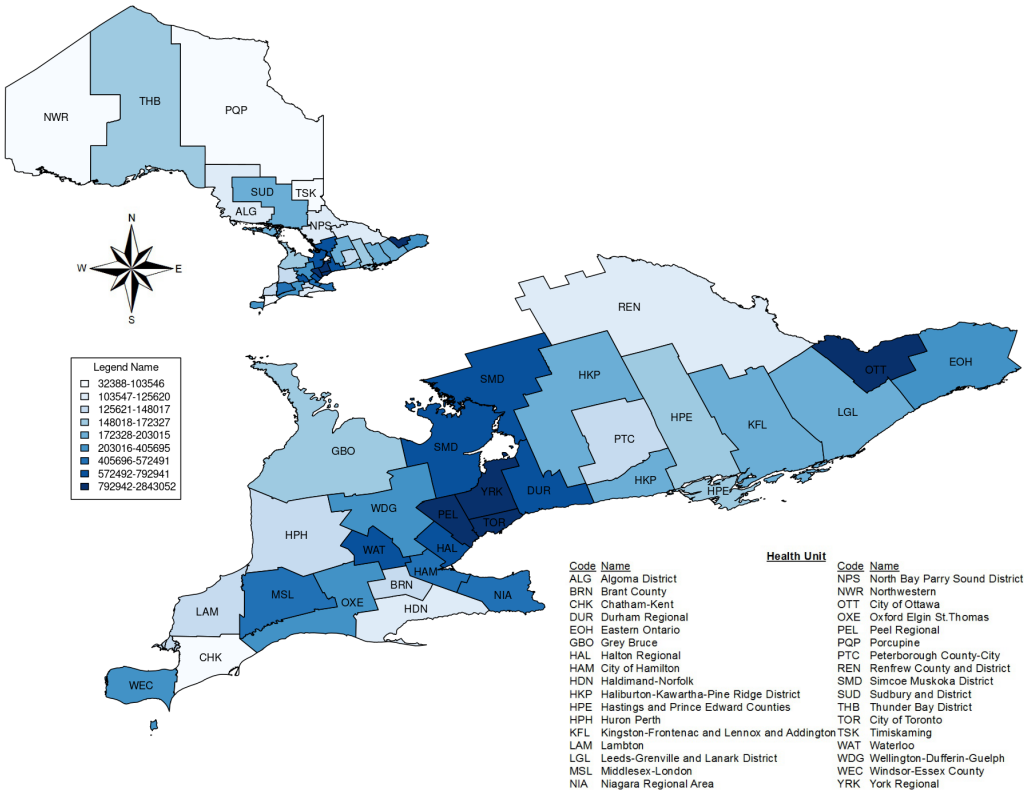


Figure 6.1: Population Map of PHU in Ontario. Created using Public Health Ontario's Easy Maps tool [53].

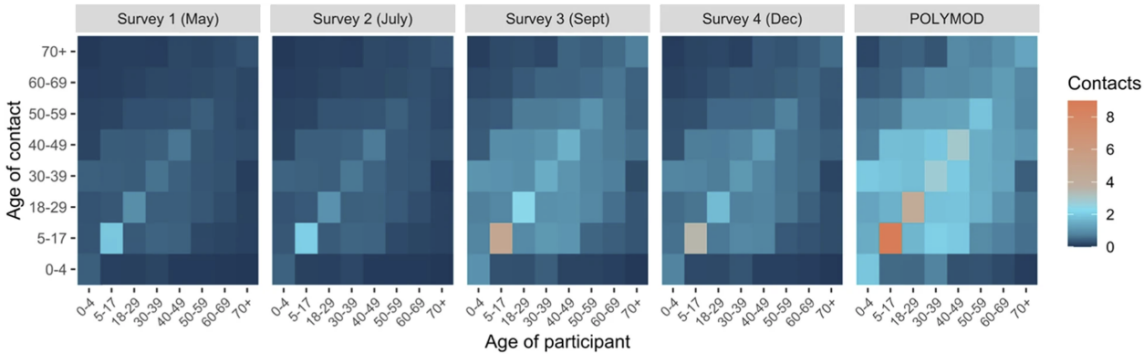


Figure 6.2: Daily Contacts by Age in Canada in 2020 (Source: Brankston et al. [11])

several groups were combined to get the number of groups down to 7: 5-17, 18-29, 30-39, 40-49, 50-59, 60-69, and 70+. This was done both to simplify the case study and to make these groups compatible with the research on contact rates between age groups in Canada during the pandemic published by Brankston et al. [11], which can be seen in Figure 6.2, which serves as the basis for the contact matrix. For the case study, we use the December 2020 survey data, as this is the survey taken closest to when the case study period begins. However, the study only provides estimates for daily contacts, so the daily contacts are multiplied by 7 to provide weekly contacts for our case study. Figure 6.3 shows the weekly contact matrix used for the case study. The values for the 0-4 and 5-17 age brackets in the original study were stated to only be estimates, as there was no data available for those age groups. This age group has by far the highest daily/weekly contact rate, which would potentially make them a good candidate for spreading the disease.

The initial conditions of the case study are based on the week of January 1st, 2021, which is when vaccines started to be administered in Ontario. Ontario has publicly available information on the number of daily active (infectious) and archived (removed) cases per PHU, which can then be distributed proportionally across the age groups to give the initial infected and removed populations. One factor that should be considered in how the case study handles the removed populations is that

Risk Group	5-17	18-29	30-39	40-49	50-59	60-69	70+
5-17	21.95	8.78	11.71	8.78	2.93	2.93	1.46
18-29	5.5	9.17	7.34	5.5	5.5	1.83	1.83
30-39	6.31	6.31	7.89	4.74	3.16	3.16	0
40-49	4.13	4.13	4.13	8.25	2.75	2.75	1.38
50-59	1.46	4.39	2.93	2.93	8.78	4.39	4.39
60-69	1.06	1.06	2.13	2.13	3.19	6.38	5.32
70+	0.32	0.63	0	0.63	1.89	3.15	5.99

Figure 6.3: Case Study Weekly Contact Matrix with Heat Map

by January 1st, 2021, many had already caught and recovered from COVID-19, which for our purposes would mean they are in the removed category. This means that once a person has contracted COVID-19, they are no longer eligible to get vaccinated under this model when in reality they would still want to eventually get vaccinated. However, the purpose of the model is to decide how to best use limited resources over a shorter time frame to save lives and minimize infections. Once a person has contracted COVID-19, vaccination becomes less of a priority, as they will have gained a natural immunity, at least for a while. With that in mind, it was decided to include any recovered cases prior to Jan 1st, 2021 as our initial removed population in the model. The initial exposed numbers are calculated as twice the initial active cases, and the susceptible numbers can be calculated as the total population of the group in the PHU minus the infectious, exposed, and recovered populations. A map of the distribution of initial cases can be seen in figure 6.4. The number of infected is likely only a low estimate of the true number of infected, as many cases would have gone unreported.

Each PHU has at least one public health unit site within its borders; and according to official policy for vaccine distribution in Ontario [3], these sites are where most vaccines are kept until they are given to local clinics and vaccination centers. This

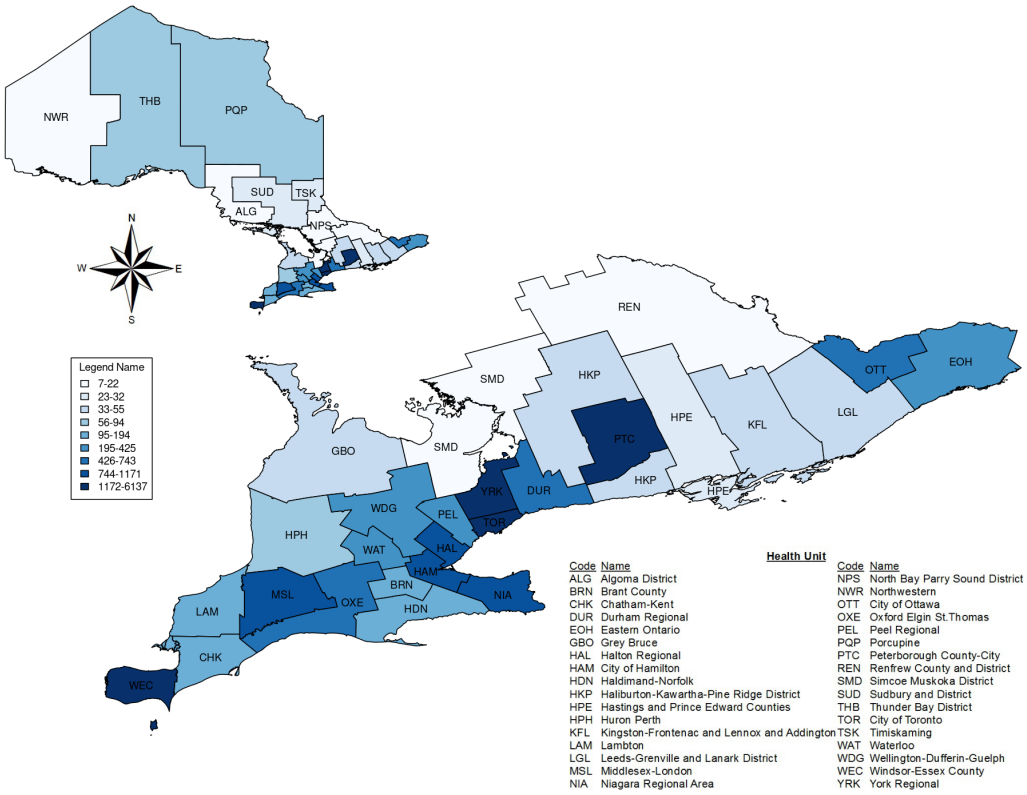


Figure 6.4: Map of Initial Active Cases. Created using Public Health Ontario's Easy Maps tool [53].

makes these public health unit sites the natural choice for the destination for the vaccines coming in from the DCs into the PHUs. In PHUs with multiple sites, the site listed as their headquarters was used. The DCs are major cold storage facilities, where vaccines are first kept when they enter the province, that are equipped to store large quantities of vaccines under safe environmentally controlled conditions for up to several weeks at a time. For the purposes of the case study, five cold chain DCs from around the province were selected as the DCs, located in Toronto, London, Ottawa, Kingston, and Sudbury. These DCs represent actual cold storage facilities that would be capable of storing vaccines and their locations were selected to provide cover over the entire province. Distances between these cold chain facilities and the PHUs were found using the driving distance between the two sites according to Google Maps, which can be found in Table 6.1. Cost per Km is $1.98 per 20,000 cubic inches of product shipped, which was based on estimates of average reefer truck shipping rates [45].

Each DC has its own maximum storage capacity, vaccine acquisition cost, and maximum volume of product it can ship in a time period, as seen in Table 6.2. Storage was loosely based on being able to store approximately 150 lots of Pfizer-BioNTech vaccine at a time, chosen as that represents roughly a month's worth of Pfizer-BioNTech vaccines. Maximum shipping volume was loosely based on being able to ship 45 lots of Pfizer-BioNTech vaccine at a time, chosen as this represents enough capacity for the DCs to collectively ship the maximum observed weekly supply of vaccines. For the case study, because there is no way of knowing where the vaccines will be coming from, to simplify the cost of getting the vaccines to the DC, each DC has a cost multiplier parameter. This represents an additional X percentage of the purchasing cost added to the costs to get the vaccines to the DC. Toronto and London being large well-connected population centers have slightly lower costs, followed by Kingstone and Ottawa. Sudbury has the highest cost multiplier, as it is more remote compared to the other locations.

Table 6.1: DC to PHU distance table

PHU	Distance (Km)				
	Toronto DC	London DC	Ottawa DC	Kingston DC	Sudbury DC
ALGOMA DISTRICT	677	838	802	922	306
BRANT COUNTY	97	85	561	378	465
CHATHAM-KENT	269	124	741	550	646
CITY OF HAMILTON	62	120	525	343	430
CITY OF OTTAWA	404	612	23	190	481
DURHAM REGION	78	235	394	206	371
EASTERN ONTARIO	457	606	93	186	588
GREY BRUCE	154	198	532	440	379
HALDIMAND-NORFOLK	140	93	604	422	509
HALIBURTON, KAWARTHA, PINE RIDGE	137	294	354	148	470
HALTON REGION	43	141	503	323	408
HASTINGS & PRINCE EDWARD COUNTIES	215	372	277	70	549
HURON PERTH	123	54	595	404	499
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	290	431	210	8	623
LAMBTON COUNTY	264	119	736	545	640
LEEDS, GRENVILLE AND LANARK DISTRICT	361	510	107	91	574
MIDDLESEX-LONDON	168	10	639	449	543
NIAGARA REGION	114	185	577	395	482
NORTH BAY PARRY SOUND DISTRICT	338	495	370	440	130
NORTHWESTERN	1,842	2,002	1,937	2,008	1,470
PEEL REGION	7	156	480	288	385
PETERBOROUGH COUNTY-CITY	145	306	282	172	393
PORCUPINE	683	844	730	800	299
RENFREW COUNTY AND DISTRICT	395	579	157	234	344
SIMCOE MUSKOKA DISTRICT	94	254	421	338	285
SOUTHWESTERN	175	30	647	457	552
SUDBURY AND DISTRICT	377	537	497	621	3
THUNDER BAY DISTRICT	1,374	1,535	1,470	1,619	1003
TIMISKAMING	491	653	524	594	227
TORONTO	39	185	412	273	384
WATERLOO REGION	84	107	556	366	461
WELLINGTON-DUFFERIN-GUELPH	63	113	535	344	439
WINDSOR-ESSEX COUNTY	343	197	815	608	719
YORK REGION	57	218	393	278	339

Table 6.2: DC Data Table

ID	Name	lat	lng	Storage (in^3)	Logistic Capacity (in^3)	Cost Multiplier
1	Toronto DC	43.665	-79.395	1,000,000	250,000	0.05
2	London DC	42.976	-81.278	700,000	250,000	0.05
3	Ottawa DC	45.403	-75.701	1,000,000	250,000	0.055
4	Kingston DC	44.239	-76.504	5,00,000	250,000	0.055
5	Sudbury DC	46.446	-80.995	700,000	250,000	0.075

Table 6.3: Vaccine Data Table

ID	Name	Efficacy	Dose Cost	Lot Size	Lot Vol. (in^3)
1	Pfizer	0.911	19.5	4875	5500
2	Moderna	0.941	15	2400	1000
3	AstraZeneca	0.740	4	2400	800

Sources: Efficacy: [1][2][51], Dose Cost:[30], Lot Size & Volume:[26]

In Ontario, three types of vaccines were administered to the public. In order of popularity they are Pfizer-BioNTech (63.4%), Moderna (33.1%), and AstraZeneca (3.44%) [37]. Each vaccine has its own efficacy, cost per dose, doses per lot, and volume requirement for shipping and storage, as seen in Table 6.3. The efficacy of each vaccine is based on the two-dose efficacy of the vaccine in the real world, even though this model uses a single dose. For the case study, a decision had to be made on what constitutes a lot of vaccines. The smallest possible lot would be a vial, where Pfizer-BioNTech contains 5 doses per vial, and Moderna and AstraZeneca contain 10 doses per vial. However, it is not realistic to ship individual vials around the province. One level up from vials is trays and cartons, where Pfizer-BioNTech groups a vial into round trays of 195, while Moderna and AstraZeneca both group a vial into cartons of 100. Then finally, Pfizer puts 25 of these trays together to make a proper shipping unit, while AstraZeneca puts 24 cartoons together into a single pallet. Moderna has its own cold storage device to store multiple of their cartons but it is unclear how many fit in a single container, it is assumed to be 24, the same as AstraZeneca. For the case study, this largest shipping size is what is considered one lot [26].

The supply of vaccines represents the number of doses of each type of vaccine that are made available each week. Ontario publishes the vaccines administered per day, which can be summed over a seven-day period to give an estimate of the weekly supply. Assuming that the province would have been trying to vaccinate as many as they could each day, this can give a reasonable estimate as to the number of vaccines

available at any given time. How many of each type of vaccine is administered is not available, so instead it is assumed that the proportions are consistent with the previously mentioned proportions of all allocated vaccines. This gives the weekly supply in doses, which is then converted into a weekly supply in lots by rounding up to the nearest full lot. Unfortunately, due to time restrictions, the case study was only ever run using the Pfizer vaccine alone, rather than all three types of vaccines. This is because the entire problem size is multiplied by the number of vaccine types. This is particularly harsh for the GMB heuristic, as it considers each vaccine type in isolation, and would take at least three times as long to run for three vaccine types compared to one. A scenario with just one vaccine type already takes approximately 26 hours to run through all solution methods.

The transmission rate for COVID-19 is difficult to determine precisely, as asymptomatic cases make it difficult to get exact case numbers and have a lower transmission rate. This is on top of having to consider different variants, differences between groups, and what restrictions are currently in place, such as whether masks are being commonly worn or not. For this case study, it is assumed there is a single transmission rate of 3% ($\tau = 0.03$) [39, 48]. Once exposed to the disease, an individual will tend to develop symptoms in 3-7 days ($\alpha_1 = 1$). Once they have developed symptoms an individual is considered infectious for the next 10 days ($\alpha_2 = 10/7$). β was set to 0.001, which represents being willing to spend \$1000 to prevent one additional infection. For the cases where we are looking to minimize new exposures, w_{ig} is equal to 1. For the case where we are looking to minimize deaths, w_{ig} are equal to the assumed mortality rates for age group g, which starting from the youngest age bracket are 0.0695%, 0.0309%, 0.0844%, 0.161%, 0.595%, 1.93%, and 4.28%, respectively [67].

The described parameters are for the base case study and might differ when testing alternate scenarios, such as altering β to put more or less weight on costs,

changing the transmission rate, changing the infectious period, or lowering the number of weekly contacts of the youngest age bracket.

6.2 Case Study Results

The solution methods run on the case study were the MC heuristic, the GMB heuristic, and the GA with a warm start from the GMB heuristic. Their performance is compared to three alternative strategies: 1) Non-Intervention; doing nothing and letting the disease run its course, 2) corrected Pro-rata; allocating vaccines proportional to a group's relative population size, but not allocating more vaccines than can be used and not wasting additional vaccines once a group has been fully vaccinated, and 3) reverse sequential by age (RS); allocating vaccines starting from vaccinating the oldest groups first and working backward through the population. Non-Intervention provides a baseline from which any strategy can be judged. Pro-rata is a very common strategy that is used to avoid bias in vaccine allocation. Reverse sequential by age is what was used in most countries, including Canada, during the COVID-19 pandemic, as it protects the vulnerable elderly population, who were at higher risk from COVID-19.

The results can be seen in Table 6.4, containing the objective value, along with its two components, epidemiological impact and cost, how many vaccines were used, and how long the method took to run for each solution method. Tables 6.5 and 6.6 outline how vaccines were allocated by the best-performing solution method (GA+GMB) when seeking to minimize new exposures and deaths, respectively. Accompanying these tables are Figures 6.5 and 6.6 that display the allocation of vaccines amongst the PHUs under both scenarios. Tables 6.7 and 6.8 show the new exposures and deaths for each PHU and age bracket under the NI strategy, accompanied by Figures 6.7 and 6.8. Tables 6.9 and 6.10 show the new exposures and deaths for each PHU and age bracket when using the strategy developed by the GA+GMB solution

method, accompanied by Figures 6.9 and 6.11. Figures 6.10 and 6.12 depict the percentage reduction observed for new exposures and deaths, respectively, between the NI strategy and that derived by the GA+GMB solution method. Figures 6.13, 6.14, and 6.15 depict the objective value and its two components under different values of β. It should be noted that as there are no decision variables for denoting deaths directly, thus when referring to deaths the values are obtained by multiplying w_{ig} (set to the mortality rate) by the new exposures (i.e., cases).

Table 6.4: Case Study Results

		Case Study Scenario					
		Base Line	Beta = 1/1250	Beta = 1/750	Beta = 1/500	Beta = 1/250	Prevent Deaths
NI	Objective Value	1,964,798	1,964,798	1,964,798	1,964,798	1,964,798	6,800
PR*	Objective Value	1,736,571	1,729,459	1,748,425	1,772,133	1,843,257	6,601
	Epidemic Impact	1,701,009	1,701,009	1,701,009	1,701,009	1,701,009	5,890
	Costs	$ 35,562,000	$ 35,562,500	$ 35,562,000	$ 35,562,000	$ 35,562,000	$ 35,562,000
	Total Vaccines	1,735,500	1,735,500	1,735,500	1,735,500	1,735,500	1,735,500
RS	Objective Value	1,969,629	1,962,515	1,981,487	2,005,203	2,076,350	6,682
	Epidemic Impact	1,934,056	1,934,056	1,934,056	1,934,056	1,934,056	5,971
	Costs	$ 35,573,000	$ 35,573,000	$ 35,573,000	$ 35,573,000	$ 35,573,000	$ 35,573,000
	Total Vaccines	1,735,500	1,735,500	1,735,500	1,735,500	1,735,500	1,735,500
MC	Objective Value	1,531,877	1,525,319	1,539,242	1,561,882	1,601,885	5,786
	Epidemic Impact	1,496,323	1,496,875	1,493,701	1,500,562	1,498,819	5,357
	Costs	$ 35,554,000	$ 35,555,000	$ 34,155,750	$ 30,660,000	$ 25,766,500	$ 21,450,000
	Total Vaccines	1,735,500	1,735,500	1,667,250	1,496,625	1,257,750	1,045,439
	Solve Time	0:30:22	0:30:20	0:30:24	0:29:59	0:30:17	0:05:54
GMB	Objective Value	1,314,427	1,305,838	1,334,128	1,372,239	1,455,658	5,789
	Epidemic Impact	1,286,763	1,283,706	1,303,900	1,330,492	1,398,937	5,595
	Costs	$ 27,664,000	$ 27,665,000	$ 22,671,000	$ 20,873,500	$ 14,180,250	$ 9,700,000
	Total Vaccines	1,238,250	1,350,375	1,116,375	999,375	692,250	472,875
	Solve Time	14:31:26	12:26:18	13:27:03	12:03:32	12:56:50	12:48:03
GA+GMB	Objective Value	1,310,118	1,304,770	1,325,835	1,347,400	1,427,411	5,740
	Epidemic Impact	1,281,354	1,282,637	1,294,142	1,303,856	1,361,899	5,541
	Costs	$ 28,764,000	$ 27,666,250	$ 23,769,750	$ 21,772,000	$ 16,378,000	$ 9,950,000
	Total Vaccines	1,404,000	1,350,375	1,131,000	1,062,750	799,500	492,375
	Solve Time	27:57:45	17:49:36	19:52:57	19:21:55	21:52:04	13:45:59

6.3 Case Study Insights

The case study results depict a much slower spread of the disease compared to the test data from earlier. Rather than burning through the entire population within the 20-week period being considered, it is instead slowly but steadily working its way through the population. This is more consistent with the behavior of COVID-19 during the pandemic, though it is still predicting significantly higher cases than what was observed in reality. According to this model, across all 34 PHUs it would be expected there would be between 60,000-90,000 new cases each week, while in reality there were closer to 24,000 cases per week [52], though this is only confirmed positive cases. This could be corrected by adjusting either the transmission rate or contact matrix. It is likely neither perfectly mimics reality since strict policies, such as quarantining, were enforced to limit disease spread, which our model does not

Table 6.5: Vaccine Allocation using GA+GMB to prevent New Exposures

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT								0	0.00%
BRANT COUNTY	14,625							14,625	1.05%
CHATHAM-KENT	9,750							9,750	0.70%
CITY OF HAMILTON	73,832	9,043						82,875	5.94%
CITY OF OTTAWA	141,312	63						141,375	10.14%
DURHAM REGION	95,901	1,599						97,500	6.99%
EASTERN ONTARIO	28,941	309						29,250	2.10%
GREY BRUCE								0	0.00%
HALDIMAND-NORFOLK	9,749	1						9,750	0.70%
HALIBURTON, KAWARTHA, PINE RIDGE								0	0.00%
HALTON REGION	92,625							92,625	6.64%
HASTINGS & PRINCE EDWARD COUNTIES								0	0.00%
HURON PERTH	14,621	1	3					14,625	1.05%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON								0	0.00%
LAMBTON COUNTY	14,624	1						14,625	1.05%
LEEDS, GRENVILLE AND LANARK DISTRICT								0	0.00%
MIDDLESEX-LONDON	67,856	10,144						78,000	5.59%
NIAGARA REGION	58,460	4,915						63,375	4.55%
NORTH BAY PARRY SOUND DISTRICT								0	0.00%
NORTHWESTERN								0	0.00%
PEEL REGION								0	0.00%
PETERBOROUGH COUNTY-CITY								0	0.00%
PORCUPINE								0	0.00%
RENFREW COUNTY AND DISTRICT								0	0.00%
SIMCOE MUSKOKA DISTRICT								0	0.00%
SOUTHWESTERN	24,375							24,375	1.75%
SUDBURY AND DISTRICT								0	0.00%
THUNDER BAY DISTRICT	4,875							4,875	0.35%
TIMISKAMING								0	0.00%
TORONTO	310,051	16,574						326,625	23.43%
WATERLOO REGION	85,125	2,625						87,750	6.29%
WELLINGTON-DUFFERIN-GUELPH	43,872	3						43,875	3.15%
WINDSOR-ESSEX COUNTY	53,625							53,625	3.85%
YORK REGION	164,750	40,000						204,750	14.69%
Total	1,308,969	85,278	3	0	0	0	0	1,394,250	
%	93.88%	6.12%	0.00%	0.00%	0.00%	0.00%	0.00%		

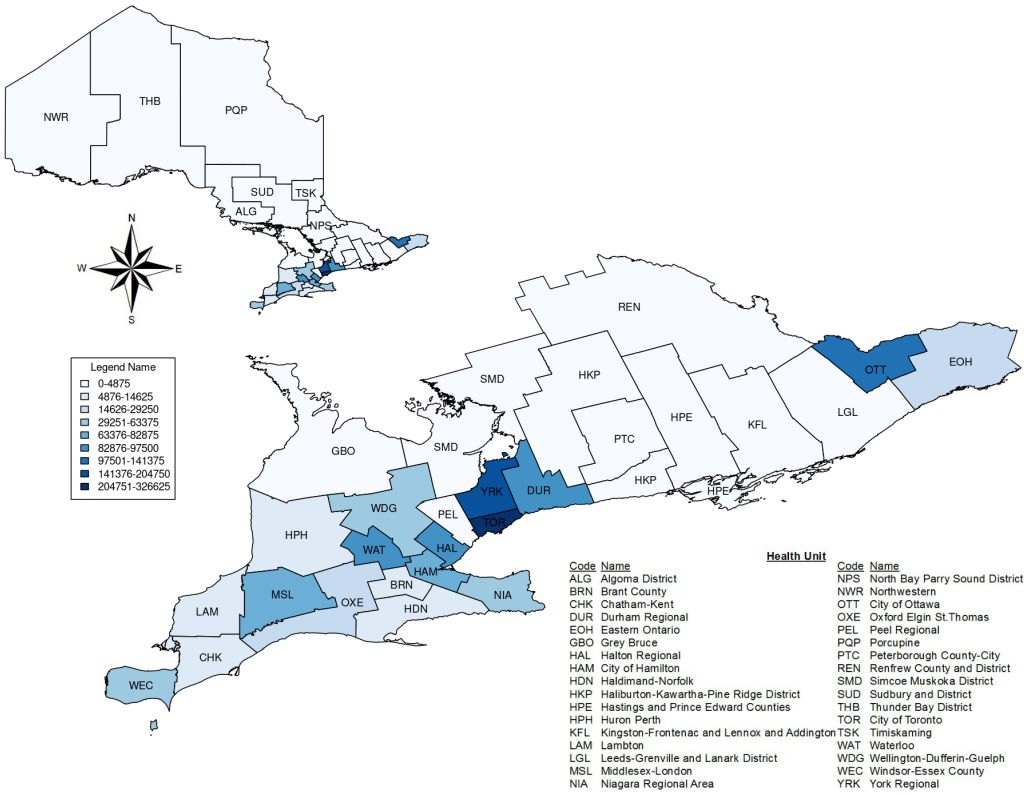


Figure 6.5: Map of Vaccine Allocation when using GA+GMB to prevent New Exposures. Created using Public Health Ontario's Easy Maps tool [53].

Table 6.6: Vaccine Allocation using GA+GMB to prevent Deaths

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT	-	-	-	-	-	-	-	-	0.00%
BRANT COUNTY	2,176	14	13	2	1,169	1,333	162	4,869	1.00%
CHATHAM-KENT	-	-	-	-	-	-	-	-	0.00%
CITY OF HAMILTON	19,500	-	-	-	-	-	-	19,500	4.00%
CITY OF OTTAWA	-	-	-	-	-	-	-	-	0.00%
DURHAM REGION	-	-	-	-	-	-	-	-	0.00%
EASTERN ONTARIO	4,875	-	-	-	-	-	-	4,875	1.00%
GREY BRUCE	-	-	-	-	-	-	-	-	0.00%
HALDIMAND-NORFOLK	-	-	-	-	-	-	-	-	0.00%
HALIBURTON, KAWARTHA, PINE RIDGE	-	-	-	-	-	-	-	-	0.00%
HALTON REGION	-	-	-	-	74	68,436	4,611	73,121	15.00%
HASTINGS & PRINCE EDWARD COUNTIES	-	-	-	-	-	-	-	-	0.00%
HURON PERTH	-	-	-	-	-	-	-	-	0.00%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	-	-	-	-	-	-	-	-	0.00%
LAMBTON COUNTY	4,875	-	-	-	-	-	-	4,875	1.00%
LEEDS, GRENVILLE AND LANARK DISTRICT	-	-	-	-	-	-	-	-	0.00%
MIDDLESEX-LONDON	14,625	-	-	-	-	24,375	-	39,000	8.00%
NIAGARA REGION	19,500	-	-	-	-	-	-	19,500	4.00%
NORTH BAY PARRY SOUND DISTRICT	-	-	-	-	-	-	-	-	0.00%
NORTHWESTERN	-	-	-	-	-	-	-	-	0.00%
PEEL REGION	-	-	-	-	-	-	-	-	0.00%
PETERBOROUGH COUNTY-CITY	-	-	-	-	-	4,534	318	4,852	1.00%
PORCUPINE	-	-	-	-	-	-	-	-	0.00%
RENFREW COUNTY AND DISTRICT	-	-	-	-	-	-	-	-	0.00%
SIMCOE MUSKOKA DISTRICT	-	-	-	-	-	-	-	-	0.00%
SOUTHWESTERN	9,750	-	-	-	-	4,875	-	14,625	3.00%
SUDBURY AND DISTRICT	-	-	-	-	-	-	-	-	0.00%
THUNDER BAY DISTRICT	-	-	-	-	-	-	-	-	0.00%
TIMISKAMING	-	-	-	-	-	-	-	-	0.00%
TORONTO	97,321	-	-	-	16	140	-	97,477	20.00%
WATERLOO REGION	-	-	-	-	-	-	-	-	0.00%
WELLINGTON-DUFFERIN-GUELPH	4,875	-	-	-	-	14,625	-	19,500	4.00%
WINDSOR-ESSEX COUNTY	24,375	-	-	-	-	19,500	-	43,875	9.00%
YORK REGION	19,500	-	-	-	-	121,875	-	141,375	29.00%
Total	221,372	14	13	2	1,259	259,693	5,091	487,444	
%	45.41%	0.00%	0.00%	0.00%	0.26%	53.28%	1.04%		

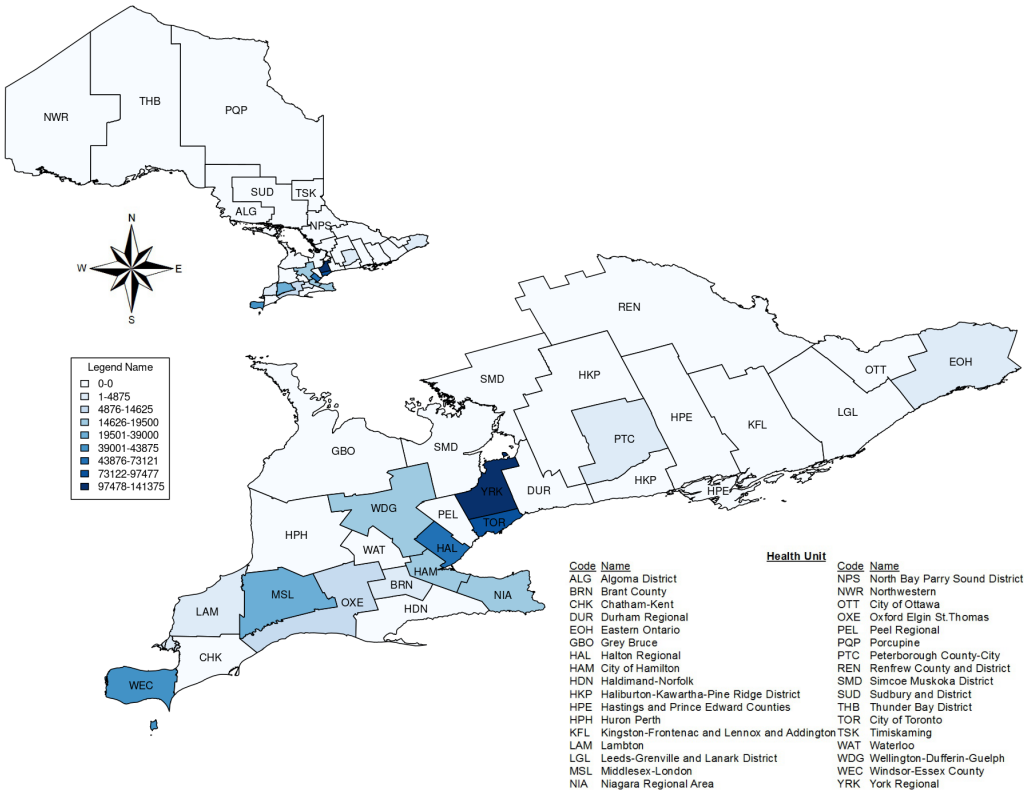


Figure 6.6: Map of Vaccine Allocation when using GA+GMB to prevent Deaths. Created using Public Health Ontario's Easy Maps tool [53].

Table 6.7: Case Study New Exposures for NI

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT	454	233	176	129	133	94	44	1,263	0.06%
BRANT COUNTY	7,046	3,701	2,933	2,155	1,850	1,079	457	19,220	0.99%
CHATHAM-KENT	4,566	2,298	1,804	1,366	1,393	916	429	12,771	0.66%
CITY OF HAMILTON	34,273	22,971	18,929	12,784	11,529	6,564	2,983	110,031	5.66%
CITY OF OTTAWA	34,469	22,479	17,234	11,864	9,730	5,072	2,174	103,022	5.30%
DURHAM REGION	32,999	17,838	14,399	10,828	9,238	4,738	1,845	91,886	4.72%
EASTERN ONTARIO	9,961	4,684	4,011	3,097	3,249	2,028	884	27,913	1.43%
GREY BRUCE	2,314	1,034	866	608	650	468	215	6,156	0.32%
HALDIMAND-NORFOLK	5,019	2,617	2,029	1,489	1,584	1,070	464	14,271	0.73%
HALIBURTON, KAWARTHA, PINE RIDGE	2,366	1,180	928	713	853	617	295	6,952	0.36%
HALTON REGION	40,165	19,652	14,853	14,153	11,063	5,111	2,420	107,417	5.52%
HASTINGS & PRINCE EDWARD COUNTIES	1,424	688	564	413	433	285	130	3,938	0.20%
HURON PERTH	4,893	2,372	1,785	1,320	1,288	853	393	12,905	0.66%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	2,924	1,869	1,402	944	894	545	257	8,835	0.45%
LAMBTON COUNTY	7,012	3,938	2,976	2,244	2,248	1,651	761	20,829	1.07%
LEEDS, GRENVILLE AND LANARK DISTRICT	2,155	1,001	803	649	730	475	213	6,026	0.31%
MIDDLESEX-LONDON	30,627	21,393	15,864	10,622	9,294	5,344	2,452	95,596	4.91%
NIAGARA REGION	26,217	16,782	11,981	9,371	9,643	5,994	3,038	83,025	4.27%
NORTH BAY PARRY SOUND DISTRICT	863	451	350	259	279	182	79	2,463	0.13%
NORTHWESTERN	1,627	724	536	388	351	219	83	3,927	0.20%
PEEL REGION	22,495	14,838	10,511	6,999	5,522	2,638	1,026	64,029	3.29%
PETERBOROUGH COUNTY-CITY	7,941	6,390	4,447	3,387	4,076	3,270	1,979	31,491	1.62%
PORCUPINE	3,632	1,844	1,399	1,044	991	625	239	9,775	0.50%
RENFREW COUNTY AND DISTRICT	914	420	388	255	249	163	73	2,462	0.13%
SIMCOE MUSKOKA DISTRICT	1,412	717	585	423	395	224	95	3,851	0.20%
SOUTHWESTERN	17,107	8,808	7,272	5,670	5,315	3,210	1,516	48,898	2.51%
SUDBURY AND DISTRICT	1,578	885	648	496	462	277	118	4,464	0.23%
THUNDER BAY DISTRICT	3,231	1,909	1,377	969	929	611	250	9,276	0.48%
TIMISKAMING	1,355	683	543	420	430	302	133	3,867	0.20%
TORONTO	151,661	140,511	127,032	75,496	60,269	31,407	15,208	601,585	30.92%
WATERLOO REGION	21,749	14,739	10,557	7,073	5,635	2,888	1,205	63,847	3.28%
WELLINGTON-DUFFERIN-GUELPH	21,749	14,739	10,557	7,073	5,635	2,888	1,205	63,847	3.28%
WINDSOR-ESSEX COUNTY	35,200	25,163	15,996	13,275	13,300	7,821	3,754	114,510	5.89%
YORK REGION	59,290	38,494	28,387	24,276	22,681	7,273	4,692	185,093	9.51%
Total	600,691	418,043	334,124	232,254	202,320	106,902	51,107	1,945,441	
%	30.88%	21.49%	17.17%	11.94%	10.40%	5.50%	2.63%		

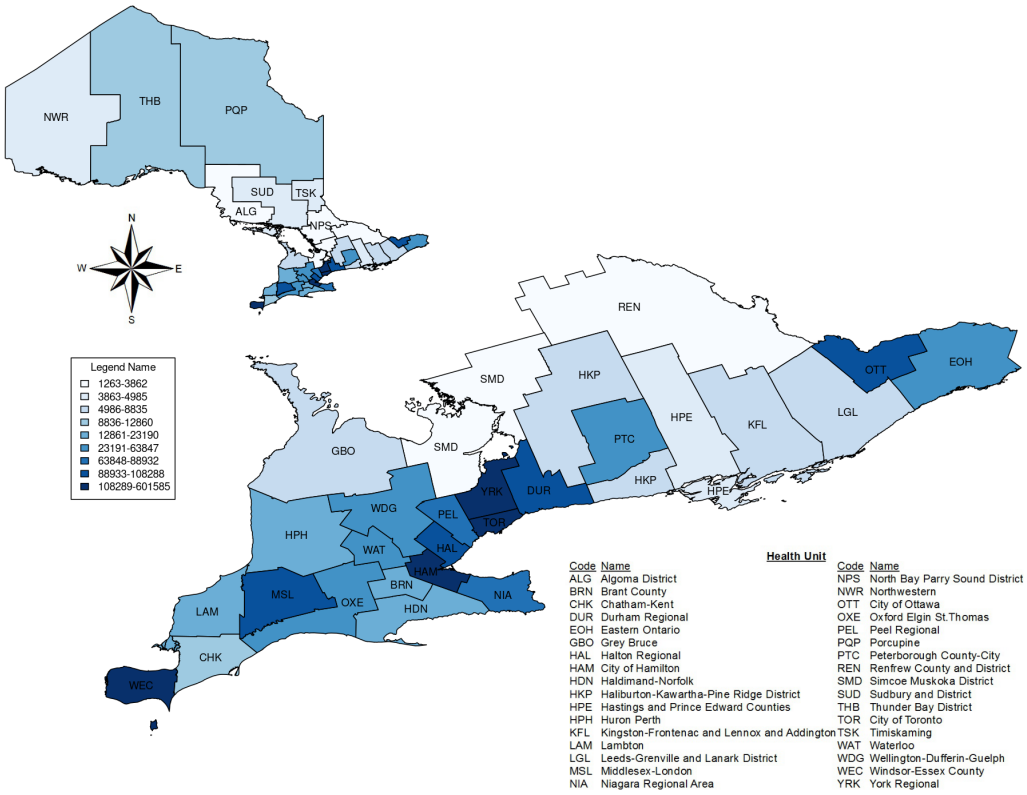


Figure 6.7: Map of New Exposures for NI. Created using Public Health Ontario's Easy Maps tool [53].

Table 6.8: Case Study Deaths for NI

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT	0	0	0	0	1	2	2	5	0.08%
BRANT COUNTY	5	1	2	3	11	21	20	63	0.95%
CHATHAM-KENT	3	1	2	2	8	18	18	52	0.78%
CITY OF HAMILTON	24	7	16	21	69	127	128	390	5.86%
CITY OF OTTAWA	24	7	15	19	58	98	93	313	4.71%
DURHAM REGION	23	6	12	17	55	91	79	283	4.26%
EASTERN ONTARIO	7	1	3	5	19	39	38	113	1.70%
GREY BRUCE	2	0	1	1	4	9	9	26	0.39%
HALDIMAND-NORFOLK	3	1	2	2	9	21	20	58	0.88%
HALIBURTON, KAWARTHA, PINE RIDGE	2	0	1	1	5	12	13	34	0.50%
HALTON REGION	28	6	13	23	66	99	104	337	5.07%
HASTINGS & PRINCE EDWARD COUNTIES	1	0	0	1	3	5	6	16	0.24%
HURON PERTH	3	1	2	2	8	16	17	49	0.73%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	2	1	1	2	5	11	11	32	0.48%
LAMBTON COUNTY	5	1	3	4	13	32	33	90	1.35%
LEEDS, GRENVILLE AND LANARK DISTRICT	1	0	1	1	4	9	9	26	0.39%
MIDDLESEX-LONDON	21	7	13	17	55	103	105	322	4.83%
NIAGARA REGION	18	5	10	15	57	116	130	352	5.28%
NORTH BAY PARRY SOUND DISTRICT	1	0	0	0	2	4	3	10	0.15%
NORTHWESTERN	1	0	0	1	2	4	4	12	0.18%
PEEL REGION	16	5	9	11	33	51	44	168	2.52%
PETERBOROUGH COUNTY-CITY	6	2	4	5	24	63	85	189	2.84%
PORCUPINE	3	1	1	2	6	12	10	34	0.51%
RENFREW COUNTY AND DISTRICT	1	0	0	0	1	3	3	9	0.14%
SIMCOE MUSKOKA DISTRICT	1	0	0	1	2	4	4	13	0.20%
SOUTHWESTERN	12	3	6	9	32	62	65	188	2.83%
SUDBURY AND DISTRICT	1	0	1	1	3	5	5	16	0.24%
THUNDER BAY DISTRICT	2	1	1	2	6	12	11	34	0.50%
TIMISKAMING	1	0	0	1	3	6	6	16	0.25%
TORONTO	105	43	107	122	359	606	651	1,993	29.94%
WATERLOO REGION	15	5	9	11	34	56	52	181	2.72%
WELLINGTON-DUFFERIN-GUELPH	15	5	9	11	34	56	52	181	2.72%
WINDSOR-ESSEX COUNTY	24	8	14	21	79	151	161	458	6.88%
YORK REGION	41	12	24	39	135	140	201	592	8.90%
Total	417	129	282	374	1,204	2,063	2,187	6,657	
%	6.27%	1.94%	4.24%	5.62%	18.08%	30.99%	32.86%		

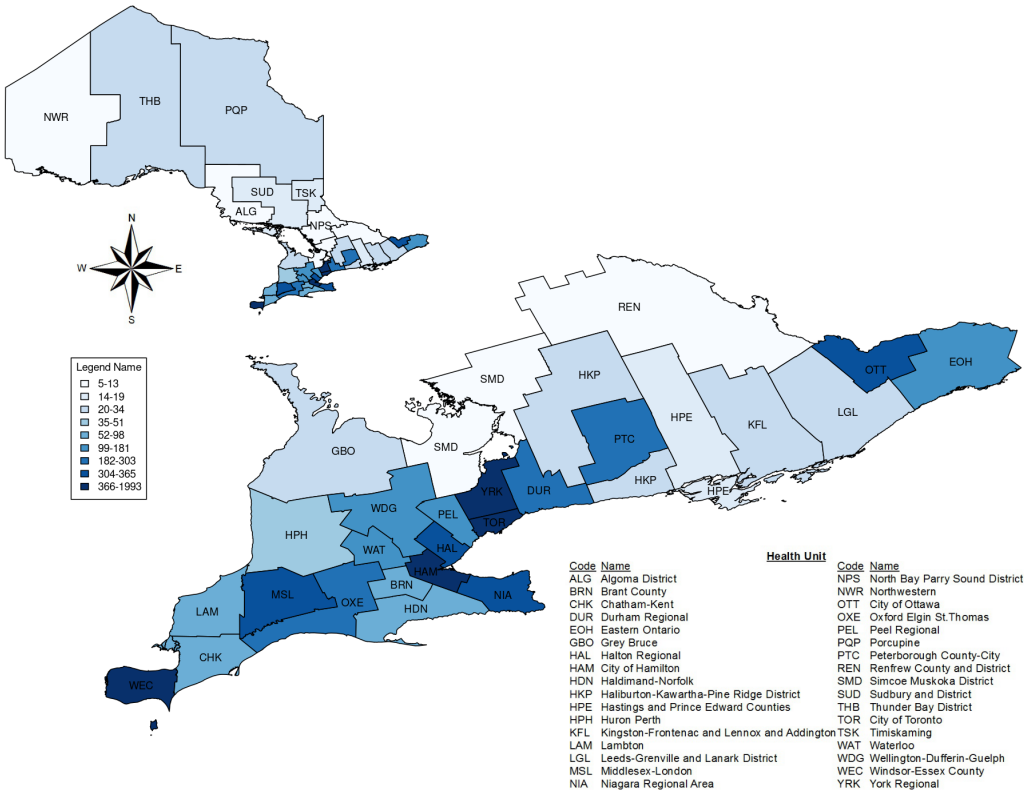


Figure 6.8: Map of Deaths for NI. Created using Public Health Ontario's Easy Maps tool [53].

Table 6.9: Case Study New Exposures using GA+GMB

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT	454	233	176	129	133	94	44	1263	0.10%
BRANT COUNTY	3,647	3,112	2,404	1,797	1,645	950	413	13,968	1.09%
CHATHAM-KENT	2,162	1,835	1,400	1,079	1,184	770	372	8,803	0.69%
CITY OF HAMILTON	11,745	16,031	12,887	8,891	8,816	4,950	2,348	65,668	5.12%
CITY OF OTTAWA	17,600	20,028	15,030	10,503	9,084	4,696	2,052	78,992	6.16%
DURHAM REGION	15,011	14,981	11,812	9,047	8,260	4,195	1,677	64,983	5.07%
EASTERN ONTARIO	4,462	3,811	3,180	2,499	2,812	1,738	780	19,282	1.50%
GREY BRUCE	2,314	1,034	866	608	650	468	215	6,156	0.48%
HALDIMAND-NORFOLK	2,438	2,104	1,585	1,185	1,355	906	405	9,978	0.78%
HALIBURTON, KAWARTHA, PINE RIDGE	2,366	1,180	928	713	853	617	295	6,952	0.54%
HALTON REGION	16,226	15,427	11,272	10,961	9,303	4,246	2,081	69,515	5.43%
HASTINGS & PRINCE EDWARD COUNTIES	1424	688	564	413	433	285	130	3,938	0.31%
HURON PERTH	3,007	2,167	1,604	1,200	1,220	803	376	10,376	0.81%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	2,924	1,869	1,402	944	894	545	257	8,835	0.69%
LAMBTON COUNTY	2,346	2,701	1,959	1,512	1,671	1,211	583	11,984	0.94%
LEEDS, GRENVILLE AND LANARK DISTRICT	2,155	1,001	803	649	730	475	213	6,026	0.47%
MIDDLESEX-LONDON	10,793	15,144	11,003	7,524	7,221	4,096	1,958	57,739	4.51%
NIAGARA REGION	9,143	11,908	8,268	6,606	7,461	4,575	2,417	50,376	3.93%
NORTH BAY PARRY SOUND DISTRICT	863	451	350	259	279	182	79	2,463	0.19%
NORTHWESTERN	1,627	724	536	388	351	219	83	3,927	0.31%
PEEL REGION	22,495	14,838	10,511	6,999	5,522	2,638	1,026	64,029	5.00%
PETERBOROUGH COUNTY-CITY	7,941	6,390	4,447	3,387	4,076	3,270	1979	31,491	2.46%
PORCUPINE	3,632	1,844	1,399	1,044	991	625	239	9,775	0.76%
RENFREW COUNTY AND DISTRICT	914	420	388	255	249	163	73	2,462	0.19%
SIMCOE MUSKOKA DISTRICT	1,412	717	585	423	395	224	95	3,851	0.30%
SOUTHWESTERN	6,310	6,307	5,005	3,978	4,088	2,434	1,197	29,319	2.29%
SUDBURY AND DISTRICT	1,578	885	648	496	462	277	118	4,464	0.35%
THUNDER BAY DISTRICT	2,793	1,876	1,349	951	921	605	248	8,743	0.68%
TIMISKAMING	1,355	683	543	420	430	302	133	3,867	0.30%
TORONTO	44,694	91,735	79,617	48,415	43,000	22,070	11,226	340,755	26.59%
WATERLOO REGION	11,028	12,979	9,139	6,216	5,226	2,657	1,131	48,376	3.78%
WELLINGTON-DUFFERIN-GUELPH	5,650	7,280	5,449	4,038	4,023	2,114	951	29,503	2.30%
WINDSOR-ESSEX COUNTY	11,233	17,440	10,604	8,960	9,902	5,721	2,865	66,725	5.21%
YORK REGION	27,678	31,110	23,167	20,172	20,292	9,940	4,411	136,770	10.67%
Total	261,422	310,930	240,882	172,663	163,932	89,058	42,468	1,281,355	
%	20.40%	24.27%	18.80%	13.48%	12.79%	6.95%	3.31%		

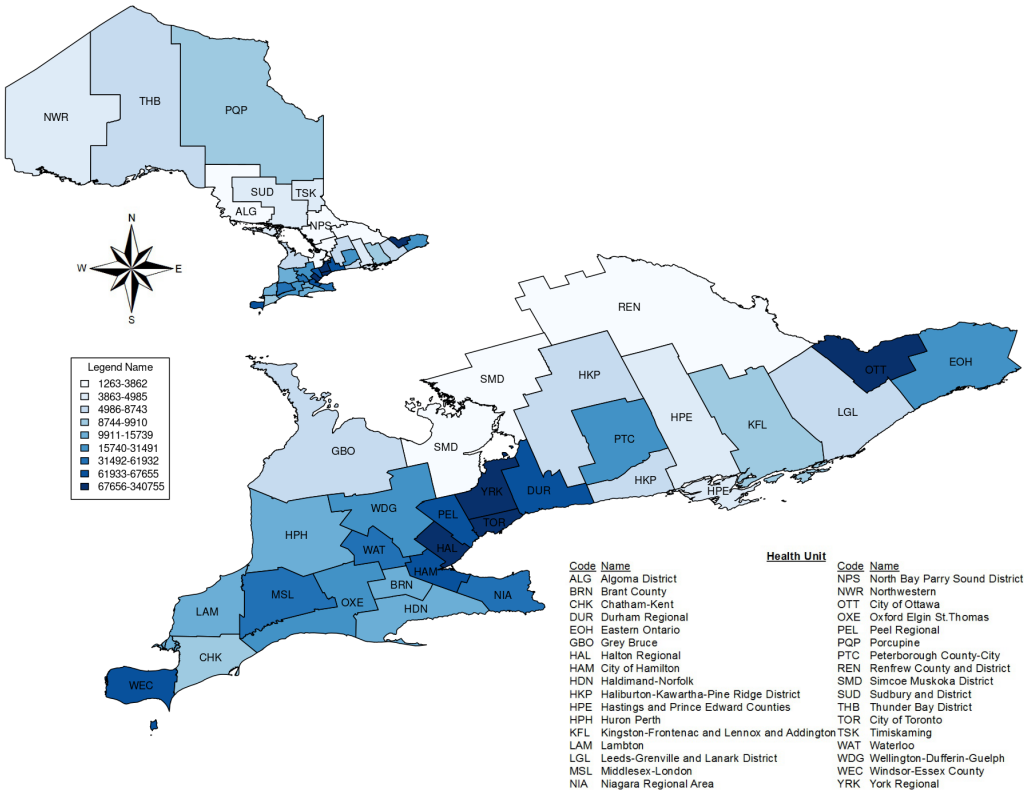


Figure 6.9: Map of New Exposures when using GA+GMB to prevent New Exposures. Created using Public Health Ontario's Easy Maps tool [53].

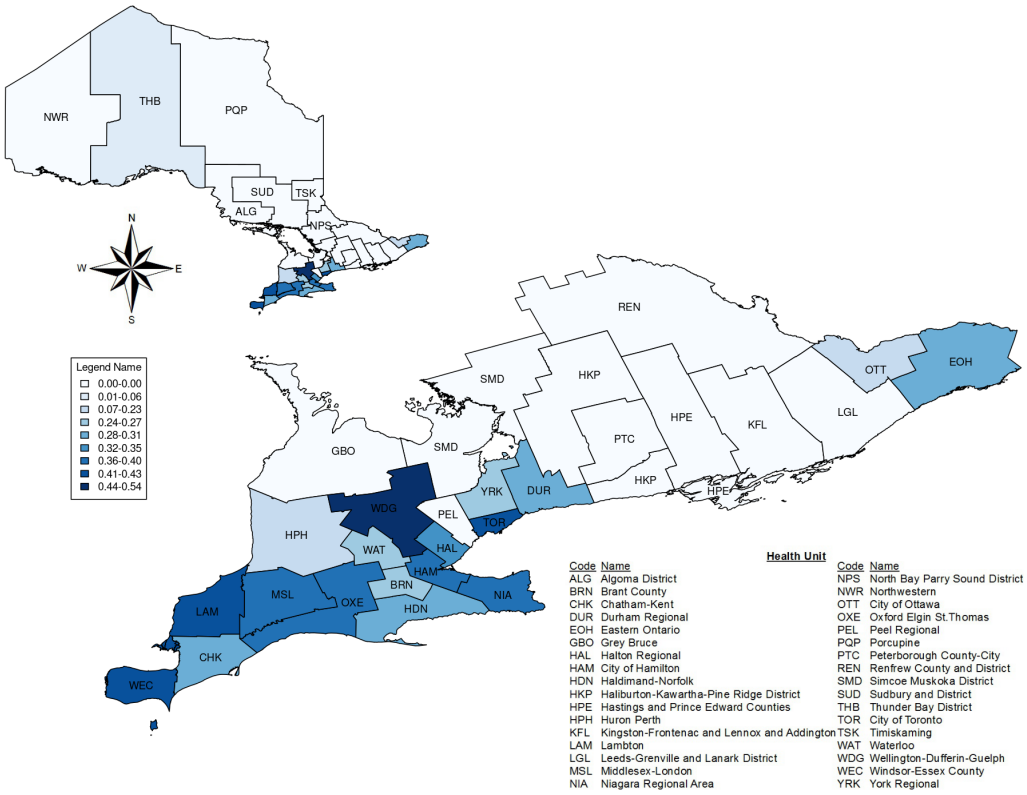


Figure 6.10: Map of the percent reduction in New Exposures using GA+GMB strategy compared to NI. Created using Public Health Ontario's Easy Maps tool [53].

Table 6.10: Case Study Deaths using GA+GMB

	5-17	18-29	30-39	40-49	50-59	60-69	70+	Total	%
ALGOMA DISTRICT	0	0	0	0	1	2	2	5	0.09%
BRANT COUNTY	4	1	2	3	9	17	17	54	0.97%
CHATHAM-KENT	3	1	2	2	8	18	18	52	0.94%
CITY OF HAMILTON	14	6	12	16	57	104	107	317	5.71%
CITY OF OTTAWA	24	7	15	19	58	98	93	313	5.66%
DURHAM REGION	23	6	12	17	55	91	79	283	5.11%
EASTERN ONTARIO	5	1	3	4	17	35	34	100	1.81%
GREY BRUCE	2	0	1	1	4	9	9	26	0.46%
HALDIMAND-NORFOLK	3	1	2	2	9	21	20	58	1.05%
HALIBURTON, KAWARTHA, PINE RIDGE	2	0	1	1	5	12	13	34	0.61%
HALTON REGION	28	6	12	22	64	47	93	272	4.91%
HASTINGS & PRINCE EDWARD COUNTIES	1	0	0	1	3	5	6	16	0.29%
HURON PERTH	3	1	2	2	8	16	17	49	0.88%
KINGSTON, FRONTENAC, LENNOX & ADDINGTON	2	1	1	2	5	11	11	32	0.58%
LAMBTON COUNTY	3	1	2	3	10	24	25	67	1.21%
LEEDS, GRENVILLE AND LANARK DISTRICT	1	0	1	1	4	9	9	26	0.47%
MIDDLESEX-LONDON	14	6	11	14	48	75	91	258	4.66%
NIAGARA REGION	9	4	7	11	44	87	101	263	4.74%
NORTH BAY PARRY SOUND DISTRICT	1	0	0	0	2	4	3	10	0.18%
NORTHWESTERN	1	0	0	1	2	4	4	12	0.22%
PEEL REGION	16	5	9	11	33	51	44	168	3.03%
PETERBOROUGH COUNTY-CITY	5	2	4	5	24	45	79	163	2.94%
PORCUPINE	3	1	1	2	6	12	10	34	0.62%
RENFREW COUNTY AND DISTRICT	1	0	0	0	1	3	3	9	0.17%
SIMCOE MUSKOKA DISTRICT	1	0	0	1	2	4	4	13	0.24%
SOUTHWESTERN	7	2	5	7	25	42	51	139	2.51%
SUDBURY AND DISTRICT	1	0	1	1	3	5	5	16	0.29%
THUNDER BAY DISTRICT	2	1	1	2	6	12	11	34	0.61%
TIMISKAMING	1	0	0	1	3	6	6	16	0.30%
TORONTO	54	32	76	87	275	460	511	1,496	26.99%
WATERLOO REGION	15	5	9	11	34	56	52	181	3.26%
WELLINGTON-DUFFERIN-GUELPH	10	3	6	8	29	42	47	145	2.61%
WINDSOR-ESSEX COUNTY	11	5	9	14	57	79	113	288	5.19%
YORK REGION	41	12	24	39	135	140	201	592	10.69%
Total	311	108	231	313	1,043	1,646	1,888	5,541	
%	5.61%	1.95%	4.18%	5.65%	18.83%	29.71%	34.07%		

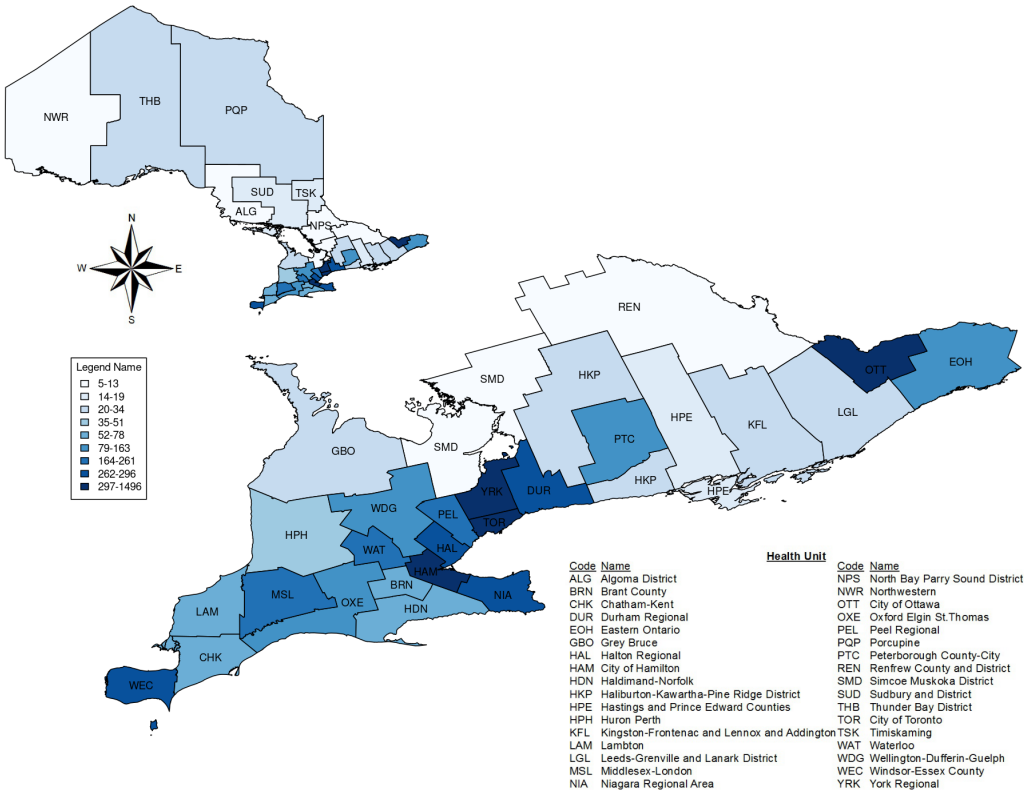


Figure 6.11: Map of Deaths when using GA+GMB to prevent Deaths. Created using Public Health Ontario's Easy Maps tool [53].

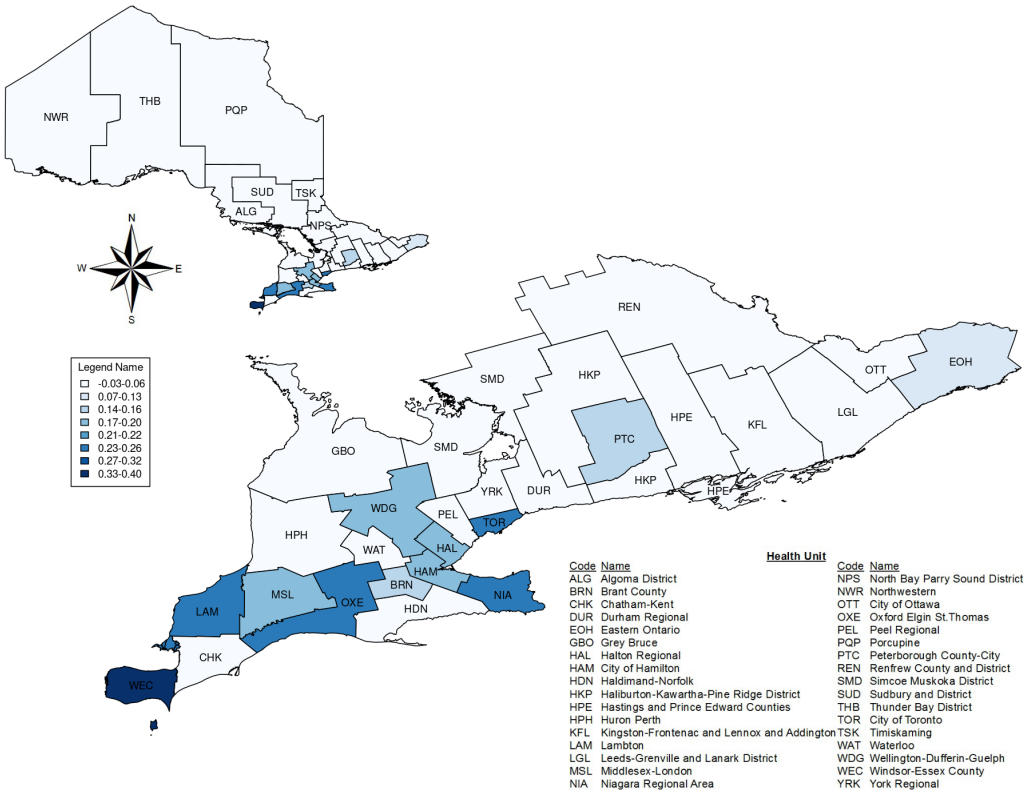


Figure 6.12: Map of the percent reduction in Deaths using GA+GMB strategy compared to NI. Created using Public Health Ontario's Easy Maps tool [53].

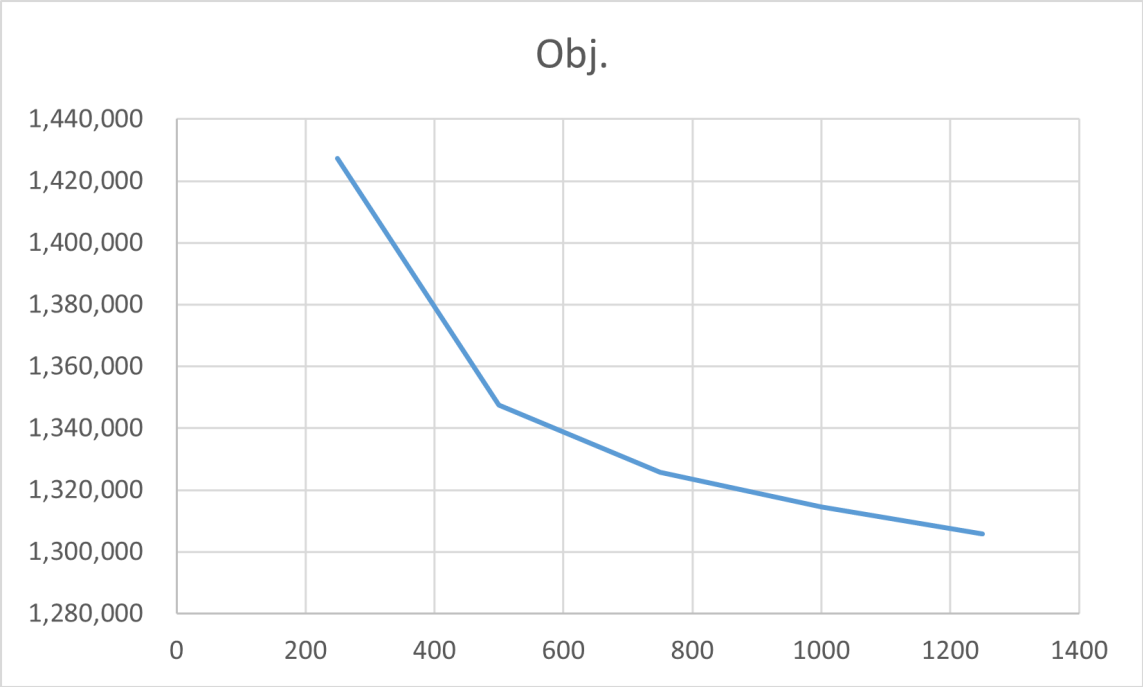


Figure 6.13: Case Study Objective Value for different values of $1/\beta$

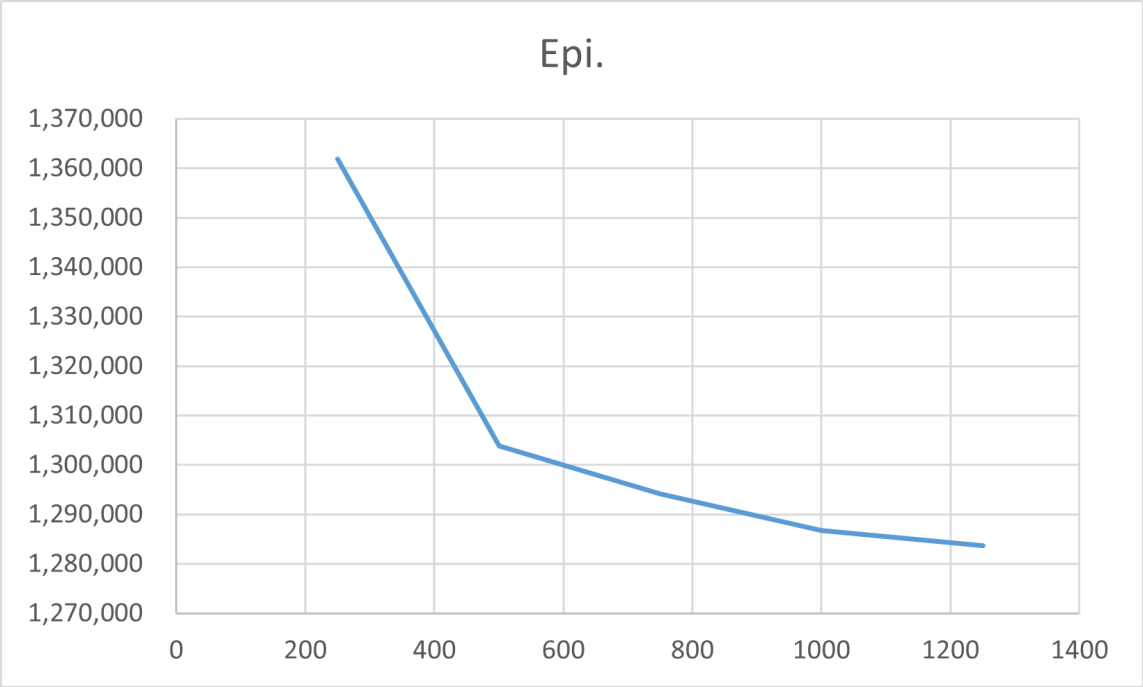


Figure 6.14: Case Study Total New Exposures for different values of $1/\beta$

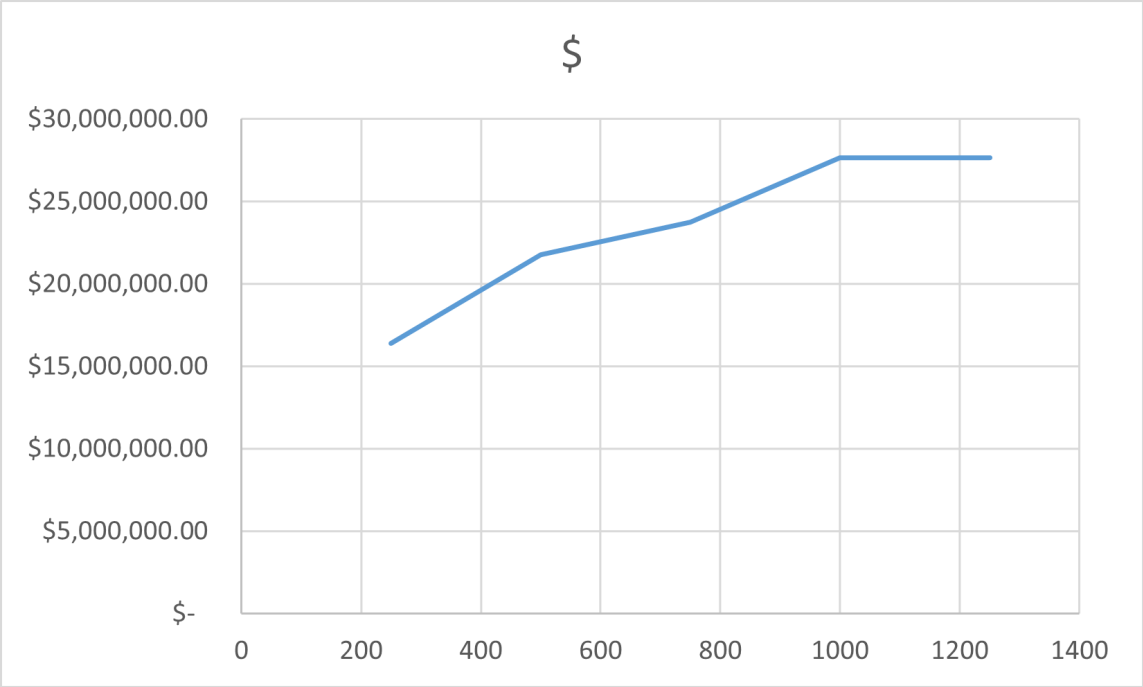


Figure 6.15: Case Study Total Costs for different values of $1/\beta$

account for, as well as there being unknown numbers of asymptomatic or unreported cases at any given time. A variable transmission rate depending on the risk group might also have been useful to consider, as some groups would have been more or less likely to transmit the disease. However, useful insights can still be gained from these results, even if they do not perfectly reflect the real world.

As can be seen in Table 6.5, the results of the case study clearly highlight the importance of vaccinating the youngest age bracket (5-18 years old) when seeking to minimize case numbers. This age bracket received over 90% of administered vaccines using the allocation strategy from the GA+GA. This is unsurprising, as according to Table 6.7, this age group had by far the most infections, accounting for 30.88% of all cases. These high case numbers are caused by the group's extremely high intra-group contact rate, meaning the entire group will quickly become infected, and thus is a large vector for spreading the disease to other groups, if not vaccinated quickly. Comparing Tables 6.9 and 6.7, total cases dropped across all age brackets, even those that did not receive a single dose of vaccines themselves. By vaccinating those who are the most likely to catch and spread the disease, not only are they themselves protected, but others are indirectly protected as well. Additionally, after implementing the GA+GA allocation strategy it can be observed that the youngest age bracket no longer accounts for the most infections, and instead, the age bracket of 18-29 makes up the majority of cases, who were beginning to be prioritized for vaccination by the end of the 20 week period. These results align with other literature on the subject [63], and the strategy of vaccinating the younger population is very common when trying to slow the spread of a disease precisely because that age range is such a potential vector for transmitting the disease. This effect would only be exasperated if schools were still open during an outbreak, as then this age bracket would have hundreds of contacts per week, as an outbreak would quickly spread through the close contact environment of a classroom.

All solution methods developed this same strategy of prioritizing vaccinating

the youngest age group in order to prevent infections. However, they did differ slightly in deciding where to vaccinate and how many vaccines to use. With the GMB heuristic and GA using fewer lots of vaccines overall compared to the MC heuristic (6.4). Additionally, even though the MC heuristic used more vaccines overall, it actually still experienced more cases than the GA+GA method, which further demonstrates the importance of where vaccines are allocated and not just how many are administered. As can be seen in Table 6.5 and Figure 6.5, only 19 of the 34 PHUs received a single lot of vaccines during the 20-week period. Even amongst these 19, 48% of the entire vaccine supply was allocated to just three PHUs, Toronto, York, and Ottawa. Why was this the case? As can be seen in Table 6.9 and Figure 6.9, many of these PHUs, such as Peel Regional and Peterborough County saw higher total new exposures compared to neighboring regions but did not receive a single lot of vaccines. At first, it was thought this might have been caused by the location of the DCs, as the two cheapest DCs are located in Toronto and London and perhaps it simply was not cost-effective to send vaccines further. However, this idea does not hold up under scrutiny as lots were sent to Ottawa and Thunder Bay, and when the weight of the costs in the objective function was lowered, a similar allocation pattern still emerged. The second hypothesis was that highly populated regions were prioritized. Comparing Figures 6.1 and 6.9, while some of the highly populated regions were vaccinated, others such as Peel Regional, the second most populous region after Toronto, were ignored. In fact, several of the least populated PHUs, such as Chatham-Kent and Haldimand-Norfolk received multiple lots. Similarly, comparing against Figure 6.4, highly populated regions with high initial case counts were ignored in favor of vaccinating other PHUs. It is likely there is no single deciding factor but multiple different interconnecting factors in play for this discrepancy in allocation, especially for the strange aversion to vaccinating Peel Regional.

When looking to prevent deaths, the results change significantly, with the different

solution methods actually producing two different strategies for solving the problem. The MC heuristic focused heavily on vaccinating Toronto (Toronto received over 66% of allocated vaccines) and focused exclusively on vaccinating the two oldest population brackets. Meanwhile, the GMB heuristic combined with the GA used fewer vaccines overall while spreading them across more PHUs. The two PHUs that received the most vaccines were York (29%) and Toronto (20%) with the remaining 51% being spread over 11 other PHUs. Additionally, the GMB heuristic did not exclusively focus on vaccinating the two oldest brackets and instead would vaccinate both the youngest (5-17) and second oldest (60-69) age brackets, using 45.41% and 53.28% of vaccines on the 5-17 and 60-69 brackets, respectively. The 60-69 and 70+-year-old age brackets are both at an elevated risk from COVID-19, but the 60-69-year-old bracket makes up a larger percentage of the population, thus vaccinating them prevents more deaths overall. Additionally, the indirect effect of vaccinating the 5-17 age bracket prevented enough additional infections to bring down the deaths in the older populations. As can be seen in Table 6.6, in Toronto where the younger bracket was especially prioritized, deaths amongst the two oldest age brackets fell from 606 and 651 (Table 6.8) to 460 and 511, respectively (Table 6.10). However, it should be noted, that while the GMB heuristic with the GA did outperform the MC heuristic in terms of overall objective value if comparing the number of deaths, it can be seen that the MC heuristic's strategy saved more lives, but incurred additional costs for the larger quantity of vaccines it used compared to the other strategy. It would be tempting to say that this makes the MC heuristic a better solution, but with the specified parameter values, saving fewer lives overall is the better approach according to the objective value. Additionally, it should be noted that neither strategy used all the vaccines available to it. This suggests that for $\beta = 1/50000$, i.e., preventing one additional death is worth spending an additional \$50 000, the value-to-benefit ratio of vaccination is on the cusp of not being worth it. Indeed, when experimenting with higher values of β before settling on 1/50 000, the optimal strategy according

to all solution methods would be to send no vaccines, as the incurred cost was not seen as worth the payoff. Even at $\beta = 1/50000$, both solution methods decide to not send any vaccines for several time periods. Trying to assign a monetary value to a life saved is part of the difficulty with using deaths and costs in the same problem and why it might be beneficial to utilize an alternate model where we instead impose a vaccination budget rather than define the trade-off between the two competing objectives, as discussed in section 3.4.

The reverse sequential by age rule-of-thumb strategy closely mirrors the strategy that was employed in reality, prioritizing vaccinating the older age brackets while distributing vaccines roughly equally across the PHUs. According to the results of this case study, this would suggest that this was not the optimal policy, both for preventing cases and for preventing loss of life. What is more, the differences are fairly large, with 1,934,056 versus 1,281,354 cases and 5,971 versus 5,541 deaths over a 20-week period. However, it is worth remembering that this model is relatively simple compared to the complexities experienced in a real pandemic scenario. For starters, it predicts far more cases than are observed in reality, and this can be for any number of reasons. Perhaps the contact matrix and transmission rate are not accurate to what was experienced at the time. Restrictions, public health guidelines, and quarantining were also in effect at the time being modeled and their impact is not captured by this model. A quarantined individual would have a drastically lower contact rate compared to others, and someone following public health guidelines might be less likely to catch or transmit the disease. If these factors were taken into account it would almost certainly slow the predicted spread of the disease. Especially if quarantining was taken into effect, meaning fewer infected individuals would not be in the general population, it might impact the optimal strategy such that prioritizing the more vulnerable populations becomes the optimal strategy.

In terms of the DCs, only the Toronto and London DCs were used as they were the cheapest, close to regions that received the majority of vaccines, and had more than

sufficient capacity to handle the volume of vaccines for this 20-week period. In fact, several weeks only required one DC to operate. However, because this case study only considers the first 20 weeks of 2021, it only has to consider the very limited supply of vaccines available during that time. Shortly after this initial 20 weeks, larger quantities of vaccines started arriving in the province, with vaccine supply going from 10-40 lots per week to over 100 lots per week by week 32. This level of vaccines might have begun to require the use of additional DCs, especially if the other two vaccine types were included as well as was originally intended. This would have brought the peak vaccine supply to 330 lots in a single week, representing over one million vaccines being allocated and distributed per week. It would be interesting to see how the strategies would change if the case study was extended to include these later weeks once the vaccine supply picked up. This additional supply also would have likely allowed for many of the high-priority groups to be fully vaccinated and it would be interesting to see how the remaining vaccines would be allocated. For instance, if the entire younger age bracket of Toronto is vaccinated, is it best to continue vaccinating Toronto's other age brackets or is it better to begin vaccinating the younger age bracket of a different PHU? As it currently stands, in the case study the logistical capacity constraints are not the limiting factor on the model as there is not sufficient supply nor administrative capacity in the regions to exceed this capacity. Instead, costs are the primary limiting factor in how many vaccines get distributed. Additionally, because the model is capable of utilizing all vaccines available to it in a time period, there is never a need to utilize the storage at the DCs. If vaccines are available and a zone has the capacity to use them, it will always be preferable to use them in an earlier time period rather than storing them for later. If supply was higher and fluctuated more, potentially with weeks of high supply followed by no supply, logistical capacity, and storage would be more of a factor.

6.4 Additional Case Study Scenarios

There are numerous additional scenarios that could be explored with this case study. One that would be interesting to explore would be to lower either the weekly intra-group contacts of the youngest age bracket or significantly reduce their weight in the objective function, so they no longer dominate the allocation of vaccines. The inclusion of a variable transmission rate depending on the risk group might also have been useful to consider so that even if a certain group might catch the disease they might not be as likely to pass it on compared to another group. Similarly, in the base case study scenario, vaccine allocation is concentrated on just a few of the PHUs, but this imbalance in allocation would be undesirable in the real world. Therefore, one potential scenario to explore would be introducing some form of equity enforcing constraints on the allocation of vaccines amongst zones to see how they impact the overall performance, such as seeking to minimize (or enforcing a limit on) the maximum difference between the vaccinated portions of the population between each zone and/or group. Alternatively, we could seek to minimize the maximum difference between the number of cases relative to population size between all zones and groups. Even just enforcing a minimum vaccination level requirement for each zone and group (so long as it is feasible to do so) could be a way of encouraging the equal allocation of vaccines. However, equal allocation of available vaccines across the population should not be confused with equity in allocation. Some people are inherently more at risk from the disease due to their age or other underlying factors, as well as where they live if it is a potential hotbed for the disease. Thus, the most equitable vaccination strategy might be to ignore the lower-risk areas in favor of high-risk areas. After all, sending vaccines where they will have less impact so as to not seem to be ignoring areas for the benefit of others would just be ignoring the areas those vaccines would have gone to without said intervention for the benefit of those new areas.

It would be interesting to explore different planning horizons for the case study. The epidemic is still underway by the end of the 20 weeks the case study covers (i.e., there are still new exposures each week), so the model is only considering a portion of the entire outbreak and thus only optimizing for that time period. If a longer time period was considered, say long enough for the disease to basically die out, the optimal strategy might change from what is currently observed. The longer the time period modeled, the larger the impact each individual vaccine will have, as the effects of earlier vaccination propagate out with time, preventing additional cases. However, considering too long of a period might prove counterproductive, as there is always going to be a deviation between reality and what the model predicts, and this deviation will only grow with each additional period considered. Thus, the most useful way to utilize the model might be to consider the full duration of the outbreak when allocating vaccines, but only make allocation decisions for the next few immediate periods, not the entire modeled period. That way the model can optimize for the long term by just considering a few periods at a time when it has the most accurate information to work with. However, knowing there will be more vaccines available in the future would impact the optimal allocation as well, as if you know you will receive more vaccines next period it might impact how you utilize your current supply, so the decisions would still be impacted by the shortened allocation time frame.

Finally, it would have been useful to perform sensitivity analysis on parameters such as the transmission rate, the contact matrix, the initial conditions of the model, and the supply of vaccines, to observe how optimal allocation behavior changes under different scenarios.

Chapter 7

Conclusion

In conclusion, this thesis makes four contributions:

First, we developed an integrated SEIR epidemiological and vaccine supply chain model that is capable of handling both heterogeneous and homogeneous problems for use in allocating and distributing vaccines over multiple periods. Because the vaccine allocation and distribution models are integrated, the vaccine allocation problem is subject to the limitations of the vaccine supply chain, including supply, logistical, and administrative constraints. This model has been demonstrated to be effective at modeling disease dynamics and the effect vaccination has on the population, as well as estimating the cost of distributing vaccines. However, the vaccine allocation problem is non-convex and discrete by time, making it computationally difficult to solve (NP-hard).

There exist several possibilities for future extensions to improve the model's ability to model epidemic scenarios. One such change could be to introduce vaccinated compartments, rather than simply moving the vaccinated to the removed compartment. This vaccinated population could still catch and transmit the disease but at a lower transmission rate and a lower mortality rate. If the vaccinated still contributed to the spread of the disease, just in a reduced capacity, it would almost certainly impact the optimal vaccination strategy. With such an extension it might likewise be useful to consider incorporating multiple doses to achieve full vaccination. Another useful extension would be to add another compartment to incorporate quarantining into the model. If some portion of the infected would go to quarantine rather than mixing with the general population, the spread of the disease could be

greatly reduced. Finally, it would be interesting to allow the model to account for how policy and behavior might change over time, as well as how the disease itself might evolve over time, such as becoming more infectious but less deadly. Varying the contact matrices, transmission rates, mortality rates, quarantining rates, etc., over time could be used to represent these factors. Potentially decision variables could even be introduced to represent whether certain policies are in place or not that impact these parameters and allow the model to decide whether to activate or deactivate said policies, at a cost.

Second, in order to use the model to identify optimal vaccination strategies for a given scenario, we explored eight different solution methods. These eight methods were: 1) Lagrangian relaxation, 2) McCormick Envelopes, 3) the MC heuristic, 4) the fixed allocation linear approximation, 5) a Piecewise-Linear approximation of the bilinear terms, 6) the GMB heuristic, 7) a GA, and 8) a logic-based BD. Of these solution methods, four returned promising results. The MC heuristic is by far the fastest of the explored algorithms and works well on small problems, but it begins to break down as the problem size increases. In general, the solutions it provides are still relatively good and can outperform simple rule-of-thumb allocation strategies such as Pro-rata or reverse sequential by age, but will be outperformed by other solution methods. The GMB heuristic is slower than the MC heuristic but performs comparably in terms of solution quality on small problems and outperforms it on medium and larger problems. The GA is slow when used alone but can be combined with either of the heuristics to improve on the quality of the heuristic solutions. Even though this improvement is often relatively small, the fact that the GA cannot easily drastically improve on the solutions suggests they are at least near optimal, with experiments on small problems suggesting that the optimality gap somewhere between 5-10% or less. Finally, the logic-based BD solution method was shown to be able to solve small homogeneous problems to the guaranteed optimal. However, the number of cuts and time required to build and solve the master problem increase

exponentially with problem size, especially with an increasing supply of vaccines, making it infeasible on all but the smallest of problems.

There exists potential to further explore the solution methods outlined in this thesis. The fixed allocation heuristic did not end up yielding very promising results using our simplistic linear approximation method, however, an alternate approach to linearization may yield better results. The GMB heuristic could be improved by exploring multiple potential portion sizes per allocation rather than a single portion, though this would come at a large increase in run time. If the reason behind the MC heuristic's inconsistent solution quality could be identified and corrected its significantly improved run time compared to the other methods would make it a very appealing solution method. Additionally, while not explored in this thesis, using the MC heuristic to warm start the GA might be more beneficial than using the GMB heuristic as it can be used to generate a warm start in a fraction of the time while the GA can make up for the slight drop in solution quality. The GA can itself be improved by further tailoring the mutation and crossover methods to take advantage of the structure of the problem and what genes are known to likely occur in better-performing solutions, though caution should be taken to not introduce excessive bias into the algorithm. Finally, the logic-based BD method could be improved both with better quality cuts and a faster method of building and solving the master problem.

Third, the model and the most promising solution methods (MC heuristic, GMB heuristic, and the GA with a warm start) were applied to a novel case study using real data from the Canadian province of Ontario during the COVID-19 pandemic. They were able to outperform simple rule-based vaccination strategies such as Pro-rata and reverse sequential by population in terms of both total cases and deaths, as well as the cost to acquire and distribute the necessary vaccines.

Finally, From the application of the solution methods to the case study, we were able to derive insights into the best practices for vaccine allocation and distribution when seeking to minimize total cases and deaths. For the tested scenario, when

seeking to minimize cases, the best practice would have been to focus on vaccinating the 5-17-year-old age bracket, as they have the most contacts per week and are a super spreader of the disease, both to their own group and others. The best-identified strategy lead to a 35% reduction in cases compared to a policy of Non-Intervention and a 25% reduction in cases over a Pro-rata allocation of vaccines. When seeking to minimize deaths, the best practice is to vaccinate a combination of the 60-69-year-old and the 5-17-year-old age brackets. The 60-69 and 70+-year-old age brackets are both at an elevated risk from COVID-19, but the 60-69-year-old bracket makes up a larger percentage of the population so vaccinating them prevents more deaths overall. Meanwhile, by vaccinating the 5-17 age bracket, we can indirectly protect the rest of the population and prevent more deaths by simply preventing cases that would have occurred if we only vaccinated the oldest groups alone. The best-identified strategy lead to a 16% reduction in deaths compared to Non-Intervention and a 13% reduction in deaths compared to Pro-rata allocation. Additionally, with such limited supplies of vaccines, the model found that focusing on a few key regions would prevent more cases and deaths compared to spreading the supply across the province. However, such a policy would likely not be considered very favorable in the real world by both decision-makers and the general public, as this inequitable distribution of vaccines could be seen as favoring specific regions to the detriment of others. A future extension of this work could explore introducing some form of equity enforcing constraints into the model, while still trying to keep cases and deaths to a minimum.

Bibliography

[1] Grading of recommendations, assessment, development, and evaluation (grade): Moderna COVID-19 vaccine. Technical report, Centers for Disease Control and Prevention, dec 2020.

[2] Grading of recommendations, assessment, development, and evaluation (grade): Pfizer-biontech COVID-19 vaccine. Technical report, Centers for Disease Control and Prevention, sept 2021.

[3] Vaccine storage and handling guidelines. Technical report, Government of Ontario, 2021.

[4] Immunization agenda 2030. Technical report, The World Health Organization, 2023.

[5] Ozgur M. Araz, Alison P. Galvani, and Lauren Ancel Meyers. Geographic prioritization of distributing pandemic influenza vaccines. *Health Care Management Science*, 15(3):175–187, 5 2012.

[6] Burcu Balcik, Ecem Yucesoy, Berna Akca, Sirma Karakaya, Asena A. Gevsek, Hossein Baharmand, and Fabio Sgarbossa. A mathematical model for equitable in-country COVID-19 vaccine allocation. *International Journal of Production Research*, 60(24):7502–7526, 8 2022.

[7] Frank Ball and Owen D. Lyne. Optimal vaccination policies for stochastic epidemics among a population of households. *Mathematical biosciences*, 177-178:333–354, 5 2002.

[8] Frank Ball, Denis Mollison, and Gianpaolo Scalia-Tomba. Epidemics with two levels of mixing. *Annals of Applied Probability*, 7(1), 2 1997.

[9] Dimitris Bertsimas, Vassilis Digalakis, Alexander Jacquillat, Michael Y. Li, and Alessandro Previero. Where to locate COVID-19 mass vaccination facilities? *Naval Research Logistics*, 69(2):179–200, 6 2021.

[10] Emanuele Blasioli, Bahareh Mansouri, Srinivas Subramanya Tamvada, and Elkafi Hassini. Vaccine Allocation and Distribution: A Review with a Focus on

Quantitative Methodologies and Application to Equity, Hesitancy, and COVID-19 Pandemic. *Operations Research Forum*, 4(2), 3 2023.

[11] Gabrielle Brankston, Eric Merkley, David N. Fisman, Ashleigh R. Tuite, Zvonimir Poljak, Peter John Loewen, and Amy L. Greer. Quantifying contact patterns in response to COVID-19 public health measures in Canada. *BMC Public Health*, 21(1), 11 2021.

[12] Christian Alvin H. Buhat, Destiny Sm Lutero, Yancee H. Olave, Kemuel M. Quindala, Mary Grace P. Recreo, Dylan Antonio S. J. Talabis, Monica Torres, Jerrold M. Tubay, and Jomar F. Rabajante. Using Constrained Optimization for the Allocation of COVID-19 Vaccines in the Philippines. *Applied Health Economics and Health Policy*, 19(5):699–708, 6 2021.

[13] Sheng Chen, Bryan A. Norman, Jayant Rajgopal, Tina Marie Assi, Bruce Y. Lee, and Shawn T. Brown. A planning model for the WHO-EPI vaccine distribution network in developing countries. *Iie Transactions*, 46(8):853–865, 5 2014.

[14] Wongyeong Choi and Eunha Shim. Optimal strategies for vaccination and social distancing in a game-theoretic epidemiologic model. *Journal of Theoretical Biology*, 505:110422, 11 2020.

[15] Ozden O. Dalgic, Osman Y. Özaltın, William A Ciccotelli, and Fatih Safa Erenay. Deriving effective vaccine allocation strategies for pandemic influenza: Comparison of an agent-based simulation and a compartmental model. *PLOS ONE*, 12(2):e0172261, 2 2017.

[16] Dipankar Dasgupta and Zbigniew Michalewicz. *Evolutionary algorithms in engineering Applications*. Springer Science & Business Media, 6 2013.

[17] Kim De Boeck, Catherine Decouttere, and Nico Vandaele. Vaccine distribution chains in low- and middle-income countries: A literature review. *Omega*, 97:102097, 12 2020.

[18] Mafalda Ivo De Carvalho, David S.M. Ribeiro, and Ana Paula Barbosa-Póvoa. *Design and Planning of Sustainable Vaccine Supply Chain*. Springer Nature, 1 2019.

[19] Lotty Evertje Duijzer, Willem Van Jaarsveld, and Rommert Dekker. Literature review: The vaccine supply chain. *European Journal of Operational Research*, 268(1):174–192, 7 2018.

[20] Lotty Evertje Duijzer, Willem Van Jaarsveld, Jacco Wallinga, and Rommert Dekker. Dose-Optimal Vaccine Allocation over Multiple Populations. *Production and Operations Management*, 27(1):143–159, 10 2017.

[21] Shakiba Enayati and Osman Y. Özaltın. Optimal influenza vaccine distribution with equity. *European Journal of Operational Research*, 283(2):714–725, 6 2020.

[22] Masih Fadaki, Ahmad Abareshi, Shaghayegh Maleki Far, and Paul P. Lee. Multi-period vaccine allocation model in a pandemic: A case study of COVID-19 in Australia. *Transportation Research Part E-logistics and Transportation Review*, 161:102689, 5 2022.

[23] Marshall L. Fisher. The Lagrangian relaxation method for solving integer programming problems. *Management Science*, 50(12 supplement):1861–1871, 12 2004.

[24] Chenyi Fu, Melvyn Sim, and Minglong Zhou. Robust Epidemiological Prediction and Optimization. *Social Science Research Network*, 1 2021.

[25] Melissa Gillis, Ryley Urban, Ahmed Saif, Noreen Kamal, and Matthew Murphy. A simulation–optimization framework for optimizing response strategies to epidemics. *Operations Research Perspectives*, 8:100210, 1 2021.

[26] Michelle R. Holm and Gregory A. Poland. Critical aspects of packaging, storage, preparation, and administration of mRNA and adenovirus-vectored COVID-19 vaccines for optimal efficacy. *Vaccine*, 39(3):457–459, 1 2021.

[27] John Hooker and Greger Ottosson. Logic-based Benders decomposition. *Mathematical Programming*, 96(1):33–60, 4 2003.

[28] Sharon Hovav and Avi Herbon. Prioritizing high-risk sub-groups in a multi-manufacturer vaccine distribution program. *The International Journal of Logistics Management*, 28(2):311–331, 5 2017.

[29] Sharon Hovav and Dmitry Tsadikovich. A network flow model for inventory management and distribution of influenza vaccines through a healthcare supply chain. *Operations research for health care*, 5:49–62, 6 2015.

[30] Jennifer Kates, Cynthia Cox, and Josh Michaud. How much could covid-19 vaccines cost the u.s. after commercialization? www.kff.org/coronavirus-covid-19/issue-brief/

how-much-could-covid-19-vaccines-cost-the-u-s-after-commercialization/, mar 2023. Accessed: 2023-07-23.

[31] Matthew James Keeling and J V Ross. Optimal prophylactic vaccination in segregated populations: When can we improve on the equalising strategy? *Epidemics*, 11:7–13, 6 2015.

[32] Jr. James E. Kelley. The Cutting-Plane method for solving Convex programs. *Journal of the Society for Industrial and Applied Mathematics*, 8(4):703–712, 12 1960.

[33] William O. Kermack and A. G. McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the Royal Society of London*, 115(772):700–721, 8 1927.

[34] Stef Lemmens, Catherine Decouttere, Nico Vandaele, and Mauro Bernuzzi. A review of integrated supply chain network design models: Key issues for vaccine supply chains. *Chemical engineering research & design*, 109:366–384, 5 2016.

[35] Michael Y. Li, Hamza Tazi Bouardi, Omar Skali Lami, Thomas A Trikalinos, Nikolaos Trichakis, and Dimitris Bertsimas. Forecasting COVID-19 and Analyzing the Effect of Government Interventions. *Operations Research*, 71(1):184–201, 1 2022.

[36] Jung Ah Lim, Bryan A. Norman, and Jayant Rajgopal. Redesign of vaccine distribution networks. *International Transactions in Operational Research*, 29(1):200–225, 12 2019.

[37] Noah Little. Covid-19 tracker canada (2020). COVID19Tracker.ca. Accessed: 2023-07-23.

[38] Ming Liu and Ding Zhang. A dynamic logistics model for medical resources allocation in an epidemic control with demand forecast updating. *Journal of the Operational Research Society*, 67(6):841–852, 6 2016.

[39] Majid Mohammed Mahmood, Noor-Ul-Ain Ilyas, Muhammad Shahzeb Khan, Muhammad Naseem Hasrat, and Nicholas Richwagen. Transmission frequency of COVID-19 through pre-symptomatic and asymptomatic patients in AJK: a report of 201 cases. *Virology Journal*, 18(1), 7 2021.

[40] Vijaya Kumar Manupati, Tobias Schoenherr, Nachiappan Subramanian, Mahalingam Ramkumar, Bhanushree Soni, and Suraj Kumar Panigrahi. A multi-echelon dynamic cold chain for managing vaccine distribution. *Transportation Research Part E-logistics and Transportation Review*, 156:102542, 12 2021.

[41] Laura Matrajt, Julia Eaton, Tiffany I. Leung, and Elizabeth R. Brown. Vaccine optimization for COVID-19: Who to vaccinate first? *Science Advances*, 7(6), 2 2021.

[42] Laura Matrajt, Julia Eaton, Tiffany I. Leung, Dobromir T. Dimitrov, Joshua T. Schiffer, David C. Swan, and Holly Janes. Optimizing vaccine allocation for COVID-19 vaccines shows the potential role of single-dose vaccination. *Nature Communications*, 12(1), 6 2021.

[43] Tomomi Matsui. NP-hardness of linear multiplicative programming and related problems. *Journal of Global Optimization*, 9(2):113–119, 9 1996.

[44] Dean Matter and Linke Potgieter. Allocating epidemic response teams and vaccine deliveries by drone in generic network structures, according to expected prevented exposures. *PLOS ONE*, 16(3):e0248053, 3 2021.

[45] Method. Trucking rates per mile 2023. www.method.me/pricing-guides/trucking-rates-per-mile/. Accessed: 2023-07-23.

[46] Mehrdad Mohammadi, Milad Dehghan, Amir Pirayesh, and Alexandre Dolgui. Bi-objective optimization of a stochastic resilient vaccine distribution network in the context of the COVID-19 pandemic. *Omega*, 113:102725, 12 2022.

[47] Zindoga Mukandavire, Farai Nyabadza, Noble J. Malunguza, Diego F. Cuadros, Tinevimbo Shiri, and Godfrey Musuka. Quantifying early COVID-19 outbreak transmission in South Africa and exploring vaccine efficacy scenarios. *PLOS ONE*, 15(7):e0236003, 7 2020.

[48] M. Murugesan, Padmanaban Venkatesan, Jagadish Ramasamy, Prasanna Samuel, Rajiv Karthik, Nabaneeta Dash, and Priscilla Rupali. Transmission dynamics of COVID-19 and utility of contact tracing in risk assessment of health-care worker exposure during COVID-19 pandemic. *Indian Journal of Community Medicine*, 47(1):82, 1 2022.

[49] C. R. Ng, T.C.E. Cheng, Dmitry Tsadikovich, Eugene Levner, Amir Elalouf, and Sharon Hovav. A multi-criterion approach to optimal vaccination planning:

Method and solution. *Computers & Industrial Engineering*, 126:637–649, 12 2018.

[50] Chantal Nguyen and Jean M. Carlson. Optimizing Real-Time Vaccine Allocation in a Stochastic SIR Model. *PLOS ONE*, 11(4):e0152950, 4 2016.

[51] Government of Canada. Astrazeneca vaxzevria covid-19 vaccine. https://www.canada.ca/en/health-canada/services/drugs-health-products/covid19-industry/drugs-vaccines-treatments/vaccines/astrazeneca.html. Accessed: 2023-07-23.

[52] Government of Ontario. Dataset - ontario data catalogue. data.ontario.ca/en/dataset/. Accessed: 2023-07-23.

[53] Public Health Ontario. Easy map tool. www.publichealthontario.ca/en/Data-and-Analysis/Commonly-Used-Products/Maps/Easy-Map/Easy-Map-Full. Accessed: 2023-07-23.

[54] The World Health Organization. Supply chain and logistics and ia2030. www.who.int/teams/immunization-vaccines-and-biologicals/essential-programme-on-immunization/supply-chain, 2023. Accessed: 2023-07-23.

[55] Panos M. Pardalos and Stephen A. Vavasis. Quadratic programming with one negative eigenvalue is NP-hard. *Journal of Global Optimization*, 1(1):15–22, 1 1991.

[56] Mir Saman Pishvaee, Jafar Razmi, and Seyed Ali Torabi. An accelerated Benders decomposition algorithm for sustainable supply chain network design under uncertainty: A case study of medical needle and syringe supply chain. *Transportation Research Part E-logistics and Transportation Review*, 67:14–38, 7 2014.

[57] Nikolaos P. Rachaniotis, Thomas K. Dasaklis, and Costas P. Pappis. A deterministic resource scheduling model in epidemic control: A case study. *European Journal of Operational Research*, 216(1):225–231, 1 2012.

[58] Mehdi Tamadon Rastegar, Madjid Tavana, Afshin Meraj, and Hassan Mina. An inventory-location optimization model for equitable influenza vaccine distribution in developing countries during the COVID-19 pandemic. *Vaccine*, 39(3):495–504, 1 2021.

[59] Yingtao Ren, Fernando Ordóñez, and Shinyi Wu. Optimal resource allocation response to a smallpox outbreak. *Computers & Industrial Engineering*, 66(2):325–337, 10 2013.

[60] Nikolaos V. Sahinidis and Mohit Tawarmalani. *Convexification and global optimization in continuous and Mixed-Integer nonlinear programming: theory, algorithms, software, and applications*. 10 2002.

[61] Ahmed Saif and Samir Elhedhli. Cold supply chain design with environmental considerations: A simulation-optimization approach. *European Journal of Operational Research*, 251(1):274–287, 5 2016.

[62] Alex Savachkin and Andres Uribe. Dynamic redistribution of mitigation resources during influenza pandemics. *Socio-economic Planning Sciences*, 46(1):33–45, 3 2012.

[63] Eunha Shim. Optimal Allocation of the Limited COVID-19 Vaccine Supply in South Korea. *Journal of Clinical Medicine*, 10(4):591, 2 2021.

[64] Shanmukhi Sripada, Ayush Jain, Prasanna Ramamoorthy, and Varun Ramamohan. A decision support framework for optimal vaccine distribution across a multi-tier cold chain network. *Computers & Industrial Engineering*, page 109397, 6 2023.

[65] Xu Sun, Eugenia Ama Andoh, and Hao Yu. A simulation-based analysis for effective distribution of COVID-19 vaccines: A case study in Norway. *Transportation research interdisciplinary perspectives*, 11:100453, 9 2021.

[66] Madjid Tavana, Kannan Govindan, Arash Khalili Nasr, Mohammad Heidary, and Hassan Mina. A mathematical programming approach for equitable COVID-19 vaccine distribution in developing countries. *Annals of Operations Research*, 6 2021.

[67] Robert Verity, Lucy C Okell, Ilaria Dorigatti, Peter Winskill, Charles Whittaker, Natsuko Imai, Gina Cuomo-Dannenburg, Hayley S. Thompson, Patrick G T Walker, Han Fu, Amy Dighe, Jamie T. Griffin, Marc Baguelin, Sangeeta N. Bhatia, A Boonyasiri, Anne Cori, Zulma M. Cucunubá, Richard G. FitzJohn, Katy A. M. Gaythorpe, W Green, Arran Hamlet, Wes Hinsley, Daniel J Laydon, Gemma Nedjati-Gilani, Steven Riley, Sabine L. Van Elsland, Erik M. Volz, Haowei Wang, Y Wang, Xiaoyue Xi, Christl A. Donnelly, Azra C. Ghani, and

Neil M. Ferguson. Estimates of the severity of coronavirus disease 2019: a model-based analysis. *Lancet Infectious Diseases*, 20(6):669–677, 6 2020.

[68] Hamed Yarmand, Julie S. Ivy, Brian T. Denton, and Alun L. Lloyd. Optimal two-phase vaccine allocation to geographically different regions under uncertainty. *European Journal of Operational Research*, 233(1):208–219, 2 2014.

[69] Xuecheng Yin and I. Esra Büyüktahtakın. A multi-stage stochastic programming approach to epidemic resource allocation with equity considerations. *Health Care Management Science*, 24(3):597–622, 5 2021.